Iterative Uncertainty Reduction via Monte Carlo Simulation: a Streamlined Life Cycle Assessment Case Study

by

Christopher E. Bolin

Submitted to the School of Engineering
in partial fulfillment of the requirements for the degree of

Master of Science in Computation for Design and Optimization

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2013

Author____

School of Engineering
May 10, 2013

Certified by____

Timothy G. Gutowski
Professor, Mechanical Engineering
Thesis Supervisor

Accepted by____

Nicolas G. Hadjiconstantinou
Director, Computation for Design and Optimization

Iterative Uncertainty Reduction via Monte Carlo Simulation: a Streamlined Life Cycle Assessment Case Study

by

Christopher E. Bolin

Submitted to the School of Engineering
on May 10, 2013 in partial fulfillment of the
requirements for the degree of
Master of Science in Computation for Design and Optimization

Abstract

Life cycle assessment (LCA) is one methodology for assessing a product's impact on the environment. LCA has grown in popularity recently as consumers and governments request more information concerning the environmental consequences of goods and services. In many cases, however, carrying out a complete LCA is prohibitively expensive, demanding large investments of time and money to collect and analyze data. This thesis aims to address the complexity of LCA by highlighting important product parameters, thereby guiding data collection.

LCA streamlining is the process of reducing the necessary effort to produce acceptable analyses. Many methods of LCA streamlining are unfortunately vague and rely on engineering intuition. While they can be effective, the reduction in effort is often accompanied by a commensurate increase in the uncertainty of the results. One nascent streamlining method aims to reduce uncertainty by generating random simulations of the target product's environmental impact. In these random Monte Carlo simulations the product's attributes are varied, producing a range of impacts. Parameters that contribute significantly to the uncertainty of the overall impact are targeted for resolution. To resolve a parameter, data must be collected to more precisely define its value.

This research project performs a streamlined LCA case study in collaboration with a diesel engine manufacturer. A specific engine is selected and a complex model of its production and manufacturing energy use is created. The model, consisting of 184 parameters, is then sampled randomly to determine key parameters for resolution. Parameters are resolved progressively and the resulting decrease in uncertainty is examined. The primary metric for evaluating model uncertainty is False Error Rate (FSR), defined here as the confusion between two engines that differ in energy use by 10%. Initially the FSR is 21%, dropping to 6.1% after 20 parameters are resolved, and stabilizing at 5.8% after 39 parameters are resolved. The case study illustrates that, if properly planned, a streamlined LCA can be performed that achieves desired resolution while vastly reducing the data collection burden.

Thesis Supervisor: Timothy G. Gutowski
Title: Professor of Mechanical Engineering

Acknowledgements

My years at MIT have been a wonderful time of personal growth, and I'm honored to have shared them with some amazing individuals.

For his support, guidance, flexibility, and kindness, I'd like to thank my advisor, Professor Gutowski. I leave this chapter of my life more informed and thoughtful than I entered - I own much of that to you. Thank you so much.

For pushing me to my intellectual limits and inspiring me to learn, I'd like to thank a few of my instructors, Dr. Roy Welsch and Dr. Juan Vielma.

For gathering mountains of data, I'd like to thank Maddie Fogler; this project wouldn't have been possible without your help and patience.

For graciously agreeing to read this thesis, I'd like to thank Professor Hadjiconstantinou.

For his advice and repartee, I'd like to thank my labmate, Dr. Sahil Sahni.

For making invaluable tools that saved me countless hours, I'd like to thank the open source artisans behind R, RStudio, Python, LyX, LaTeX, BibDesk, and BibTeX, all of which contributed to this thesis.

For many conversations over breakfast and debates over beer, I'd like to thank my good friend Steve Renter. "As iron sharpens iron..."

For hours of proofreading and years of unconditional care, I'd like to thank my parents, Jim and Ruth Bolin. This thesis and degree are testaments to the values you both instilled in me. Thank you for your love and dedication.

For a closer camaraderie than I ever thought possible, I'm honored to thank my partner, Jennifer Lynn. Thank you for journeying through life with me - I cannot wait to see what's next.

. . .

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation: Easier Life Cycle Assessments

Life cycle assessment (LCA) is a tool used in both the public and private sector to assess a product or service's impact on the environment. In the private sector, LCA has been increasingly used by manufacturers and retailers to meet consumer requests for more information about the environmental impact of their products. Along with the increased use of LCA comes a demand for more effective and efficient LCA methods.

To begin an LCA, data must be collected on the environmental impact of the various components in a product. As products may be very complex, guidance is needed on which data to collect. LCA streamlining is the process of reducing the necessary effort to produce an acceptable LCA. While these methods can be effective, the reduction in effort is often accompanied by an increase in the uncertainty of the results. The area of LCA streamlining is currently very active, with many new methods being proposed and tested.

1.2 Goal: Create Streamlined Assessment Procedure for Diesel Engine Production

This research project analyzes LCA streamlining techniques as part of an ongoing collaboration with a diesel engine manufacturer. The goal of this research is to improve LCA methodology for performing accurate but streamlined LCAs of diesel engine manufacturing. A case study is undertaken to this end. The target product for the case study is a large 15-liter on-highway truck engine, a relatively complex assembly consisting of over 1,000 individual parts.

1.3 Outline

Chapter 2: Background This chapter introduces this project's central case study and industry collaboration. The life cycle assessment (LCA) methodology is also discussed, as well as motivations for LCA. The concept of "streamlining" to more efficiently create useful LCAs is addressed, along with recent examples. The Monte Carlo simulation technique is introduced, with a specific emphasis on its relation to LCA. Finally, the specific application area for this research, diesel engine production, is discussed.

Chapter 3: Modeling and Simulation Methodology This chapter specifies the basis of the modeling and simulation used in this research. The concept of reducing uncertainty iteratively by increased specification is highlighted. The details of simulation modeling are discussed. The data sources are introduced, as well as the metrics used to evaluate the simulation: False Signal Rate (FSR) and Partial Rank Correlation Coefficients (PRCC). Finally, the parameters of the engine model are enumerated.

Chapter 4: Simulation Results and Analysis This chapter presents the results of the simulations, as well as necessary changes that were made to the experiment in light of initial results. The importance of a handful of influential parameters - overall mass and composition - was underestimated by PRCC. The themes of the results are discussed. It was found that very few parts have appreciable impact on the uncertainty of the model. Additionally, the uncertainty of the model has a distinct lower bound, which was not surpassed after resolving all resolvable parameters.

Chapter 5: Additional Engine Production Assessments Separate from the impact modeling and simulation, other brief assessments of diesel engine production were also explored. This chapter discusses four such assessments: an approximate life cycle assessment of a diesel engine (including carbon emissions and water usage), an engine camshaft design comparison, an analysis of resource use at the factory scale, and an investigation into the material composition of a wide range of diesel engines.

Chapter 6: Conclusion This chapter summarizes the findings of this thesis.

Chapter 2

Background

2.1 Life Cycle Assessment

Life cycle assessment (LCA) is a methodology to evaluate the overall impact of a product or service on the environment. Individual applications of LCA may be qualitative or quantitative [1, 2]; these LCAs may focus on broad or narrow categories of impact, such as energy use, carbon dioxide emissions, global warming potential, or toxicity [3]. The applications of LCA are also tremendously diverse; for example, LCAs have been used by private organizations to evaluate the impact of their products [4] and by governments to label goods for sale [5].

2.1.1 LCA Stages

Despite the variety of life cycle assessment approaches and forms, there are some unifying themes. Almost all LCAs concentrate on multiple stages of the product's life; these stages are commonly called life cycles. While the definition of these stages varies slightly, the following general stages are featured in many LCAs (adapted from the Environmental Protection Agency [6]).

1. Raw material extraction

2. Materials production

3. Manufacturing

4. Use

5. Reuse
6. Maintenance
7. Waste management and recycling

In specific LCAs, these stages may be expanded or grouped together. Some LCAs may evaluate a subset of stages, while others aim to include the entire life cycle. All-inclusive LCAs may be "cradle-to-grave", raw material acquisition through disposal, or "cradle-to-cradle", raw material acquisition through recycling.

2.1.2 Comparative Nature of LCA

LCA is always used in the context of comparison. Rather than deriving absolute results, LCAs compare two or more items in relation to each other [7]. Below are examples of possible comparisons.

Competing products and services. Two or more products or services that fulfill the same function can be compared using LCA. A "functional unit" must be declared for the analysis. The functional unit is a quantitative measure of the common service delivered by each competing item. The impact of each product or service is then normalized to the same functional unit. As an example, in Dettling [4], paper towels are compared with two varieties of electric hand dryers. The functional unit in this study was drying 260,000 pairs of hands.

Variations of the same product. Rather than comparing different competing products with LCA, variations of the same product could be compared to evaluate the impact of key design changes. For example, vehicle lightweighting has been suggested by some [8] as a potential energy-saving option over the vehicle's lifetime; in these analyses a vehicle is compared with a hypothetical lightweight variant. This can be evaluated with an LCA or with simpler methods, as has been done by Ashby [9].

Different components of the same product. The various parts or materials that constitute a product may also be compared to each other, usually with the intent of identifying those with the most significant impact. In a thorough LCA study Sullivan et al. [10] represented a generic family sedan with over 600 constituent parts .

Different stages of the same product. As mentioned earlier, LCAs study various stages of a product's life. The environmental impact of these stages can be compared and contrasted with each other. For example, Gutowski et al. [11] compare the life cycles of various product categories to evaluate the energy savings of remanufacturing versus new production. The energy used by a remanufactured product (e.g. an automotive tire) in its use phase (e.g. driving) might offset any energy savings in the remanufacturing phase, when compared to a comparable new product.

In should be emphasized that results from separate and unrelated LCAs are very difficult to compare, as many of the underlying assumptions may be fundamentally different. This is true even if the target product or service is identical.

2.1.3 Specificity in LCA

LCAs have been compiled with varying levels of thoroughness, from qualitative observations to rigorously quantitative measurements. Many researchers have described these levels. Henrik Wenzel summarized the spectrum of LCAs in three basic levels [1]. In increasing order of complexity and specificity, they are

1. **Matrix LCA.** This level may be either quantitative or qualitative. Any calculations are very rudimentary.

2. **Screening LCA.** This LCA is quantitative and based on secondary data from existing databases. No new LCA-specific measurements are made. Calculations may be more advanced.

3. **Full LCA.** This LCA includes new application-specific measurements and data. Calculations are typically more advanced.

Increasing the specificity of an LCA will heavily affect the associated workload. It is therefore very important to define the scope and purpose of an LCA early in the process.

2.1.4 LCA Standards and Guidance

The increasing use of LCAs has prompted the creation of new standards and guidance, most noticeably those introduced by the International Organization for Standardization (ISO):

ISO 14040:2006 Life cycle assessment - Principles and framework [12]

ISO 14044:2006 Life cycle assessment - Requirements and guidelines [13]

As a testament to the ongoing growth and change in LCA, the ISO is also developing two new standards. ISO 14071 will add review processes and reviewer competency requirements; ISO 14072 adds additional requirements and guidelines specifically for organizations [14].

2.1.5 Uncertainty in LCA

The theme of uncertainty in LCA is closely related to LCA streamlining. In some assessments, streamlining is conducted by targeting activities that contribute most to the uncertainty of the impact results. The sources of uncertainty in LCA have been categorized by Eric Williams [15] as follows.

Data Uncertainty Data uncertainty is caused by both errors in the data collection process as well as the imprecision of the measurements.

Cutoff Uncertainty Cutoff uncertainty is caused by the finite boundaries of the analysis, which necessarily leave some activities unaccounted.

Aggregation Uncertainty Aggregation uncertainty arises when different processes are grouped together into larger, or general "superprocesses". This is most common in economic input-output model LCA, or EIO-LCA, which determines the impact of a product or service based on its cost.[1]

Temporal Uncertainty Temporal uncertainty is caused by the fluctuation of activity impacts over time. Over a large enough time scale, the impact of activities can change appreciably, primarily due to technological improvements and government regulations. This timescale differs for each process but may be as short as a few months.

[1]See [16] for more on EIO-LCA

Geographic Uncertainty Geographic uncertainty is caused by the geographic variation in activity impacts. As regulations and technologies differ drastically between economies, so do the environmental impacts associated with production in these economies.

2.2 Streamlined LCA

Because of the increasing use of LCA, there is a growing interest in reducing the effort and time invested in an LCA. The concept of Streamlined LCA (SLCA) was first introduced in the early 1990s by Keith Weitz et al. at the Research Triangle Institute [17]. SLCA is the general approach of conducting abridged LCAs. SLCA is not a specific strategy, but a general grouping of possible strategies to reduce effort in LCA while maintaining relevance.

Many methods of LCA streamlining are unfortunately vague and rely on engineering intuition. While they can be effective, the reduction in effort is often accompanied by a commensurate increase in the uncertainty of the results. In fact, in a study of ten different streamlined techniques Hunt [18] found that over half came to an incorrect conclusion when compared with a full LCA. In light of this, streamlining methods for LCA should be evaluated critically.

2.2.1 Primary Streamlining Strategies

The primary strategies described by Weitz [17] are summarized below.

- Create a list of damaging activities to be assessed. This is sometimes known as an "inviolates" list. As an example, the inviolates list may include toxic chemicals and substances as described by the United States Environmental Protection Agency. This strategy relies on the judgment of the researchers, and may ignore large impacts, especially previously unknown impacts.

- Remove activities of subjectively minor importance. This strategy again relies on the judgment of the researchers, and previous biases can potentially impact the assessment dramatically. This strategy mainly draws upon previous LCA work to narrow the assessment scope.

- Restrict the assessment to a subset of the life cycle stages. This strategy can be very powerful, especially when comparing two closely related products that may share similar life cycles stages. In broader applications this strategy must necessarily rely on subjective judgment and previous work.

- Cull environmental impacts. This strategy is implemented to some degree in every LCA. For example, greenhouse gas emissions are commonly one of the primary impacts studied and reported.

- Cull inventory parameters and variables. This strategy eliminates some impact assessments. For example, if greenhouse gas emissions are to be reported in an assessment, energy use for each activity must be included in the inventory.

- Select only high mass or high volume activities. This strategy relies on quantitative measures of activity size; only the cutoff point is subjective. In Hunt [18] this strategy was used to alternatively exclude materials with mass less than 10% and 30% of the total. In many situations, however, this strategy eliminates important activities. For example, rare earth metals are commonly used in relatively small quantities per product, but their impact may be exceptionally high; this is primarily due to their intensive extraction processes.

- Abbreviate or eliminate impact assessments. This strategy is very restrictive. The resulting LCA has very limited application, as it is not possible to evaluate options based on their potential environmental impact.

- Utilize qualitative assessments. This strategy is also very aggressive, and relies strongly on the subjective judgment of the researchers. The primary weakness of the qualitative assessment approach is that it is very difficult to compare and contrast different activities.

- Use surrogate data for an activity. In many LCAs this strategy is a necessity. The choice of surrogate data is highly subjective, and in many cases can significantly alter the assessment results.

2.2.2 Streamlined LCA Themes

The above strategies to streamline LCAs have a few key shared characteristics. Importantly, each streamlining method is subjective in nature. The judgment of the assessment researchers is important and influential. Because of this subjectivity, each strategy has the potential to be controversial. Researchers may remove an environmental impact that they believe will be inconsequential for an assessment; a future study may discover that this impact category was instead very important, potentially nullifying many of the previous conclusions.

Despite their faults, streamlining strategies are necessary. Graedel [19] emphasizes that every life cycle assessment will be incomplete. This incompleteness is a product of many factors. First, each assessment must have discrete, finite boundaries; the impact of the studied activities outside of these boundaries will not be captured. Second, there are limitations to the amount and quality of data measurements. For example, even if experimental data can be collected for primary processes, many of the secondary processes may rely on external data sources. Finally, geographic, process, and temporal variability all necessitate limiting the resolution of the assessment. For example, the researchers must either choose specific geographies to study, or aggregate the impacts in disparate areas. The incomplete nature of LCA is more thoroughly studied in the context of uncertainty.

2.2.3 Streamlined LCA Examples

Hunt [18] provides a detailed analysis of different specific streamlining strategies, many of them variations or specific instantiations of Weitz's methods. Hunt includes the following nine strategies:

- Removal of upstream components. Only the fabrication, use, and disposal life cycle stages are studied.

- Removal of partial upstream components. No components before material manufacturing are studied, with the exception of the step just before material manufacturing. This method is slightly more inclusive than the preceding.

- Removal of downstream components. Only processes up to and including material manufacturing are studied.

- Removal of upstream and downstream components. Only the material manufacturing phase is studied.

- Specific impact used to represent entire impact categories. Rather than the use of all components (e.g. solid waste), a representative component is selected (e.g. plastic waste).

- Qualitative or less accurate data used. For components that contributed less than 10% to the LCA results, proxy data was used.

- Surrogate processes used. Based on data availability, processes were replaced with others that were physically or chemically similar.

- Exclude materials with mass $< 10\%$ of total. Any raw material with a mass of less than 10% of the total mass of all materials was not studied.

- Exclude materials with mass $< 30\%$ of total. This is simply a more aggressive and less inclusive form of the preceding method.

The researchers conclude that the most promising technique is to use qualitative or less accurate data for components that are not significant in the overall impact. The use of proxy data was further explored by Hochschorner and Finnveden [2]; and Hur [20].

In this same theme, Patanavanich [21, 22] and Zgola [23] use less accurate data for less significant components. Notably, their approach to qualifying the significance of components differs from the other researchers. Uncertainty analysis is integrated into the assessment process from the outset; the significance of components or parameters is their contribution to the overall uncertainty of the results. This approach relies on Monte Carlo simulations, which are discussed in this context in Section 2.3.

In a novel streamlining method, Sousa [24, 25, 26] uses machine learning techniques such as artificial neural networks to approximate the impact of product concepts. This method relies on existing LCAs that train the learning model. The model then associates combinations of product attributes to environmental impacts such as energy, solid waste, and smog. It should be noted that this technique was created specifically for the early stages of product design. Even so, this method was able to successfully rank the impact of various products [24].

2.3 Monte Carlo Simulations in LCAs

As mentioned above, Monte Carlo simulations are utilized in some LCA techniques. This section briefly introduces the concept of Monte Carlo simulations, with specific emphasis on the technique's use in LCA.

2.3.1 Background on Monte Carlo Simulation

Monte Carlo simulation is a stochastic method that has been used in a variety of applications [7, 21, 23, 27]. Monte Carlo methods are based around one or more mathematical or computational models of a system. Each parameter in the system is declared with a probability distribution or a fixed value. Monte Carlo methods sample system parameters from their corresponding probability distributions; from these sampled system parameters, an output vector is calculated. This simulation is repeated, usually for hundreds or thousands of iterations. The distributions of the output vectors are then used to determine uncertainty bounds, usually at certain quantiles, such as 5% and 95%.

In LCA, the model is usually constructed to estimate the environmental impact of a product or service for certain stages of its life cycle. Examples of parameters include product attributes, material composition, use-phase duty cycles, and electricity generation details. The output vector of these LCA models consists of one or more environmental impacts, such as embodied energy or greenhouse gas emissions.

2.3.2 Simulation to Quantify Uncertainty

Many LCA studies provide uncertainty ranges for final impacts. These ranges are usually produced with uncertainty analyses that rely on simulation techniques like Monte Carlo simulation. Maurice [7] and Sonnemann et al. [27] use Monte Carlo simulation to analyze uncertainty in their life cycle assessments, which respectively target coal power plants and waste incinerators. Sonnemann et al. assign various uncertainty distributions to their model parameters based both on extensively available data and on expert opinions, depending on availability. Monte Carlo simulations are then executed to provide probabilistic distributions of the impacts rather than solitary concrete values. In their study they claim that these distributions "correspond to a better understanding of the magnitude of the uncertainties in LCA results." [27]

2.3.3 Simulation to Streamline LCA

One method to streamline the LCA process is uncertainty reduction. Uncertainty reduction is typically an iterative process. Initially, most product parameters are left very general with large relative variances, which represents high uncertainties. Initial LCA impact calculations, therefore, have an unacceptably high level of uncertainty. Iteratively, product parameters are "resolved," or set to specific values with low variance; the parameters with the highest contribution to uncertainty are selected for resolution. The impact metrics and their associated uncertainties are calculated, and the process continues until the uncertainty is within pre-determined bounds. This procedure can be carried out before data collection to create a comprehensive list of data that must be gathered.

Figure 2.1 on the following page illustrates the process of uncertainty reduction to assist in differentiating between two products.

Examples of this method include the above mentioned work by Patanavanich [21, 22] and Zgola [23]. In Zgola, liquid crystal displays manufactured for laptop computers were studied. The parameters of these products included the product's lifetime, total mass, the number of LCD bulbs, the bulb type, the screen size, the laptop's duty cycle, and the electrical grid's fuel mixture[2]. In this study, the attributes that contributed most to the uncertainty were (in decreasing order of contribution) product duty cycle, electrical grid's fuel mixture, and product lifetime. The study's target error rate was 10%; this resolution was achieved after 22 parameters were resolved from a total pool of 47 total product parameters.

Patanavanich uses Monte Carlo simulation in a similar way, with a specific focus on the underspecification of material properties. In this work, all materials are described with five "levels" of increasing specification: material category, material property, material type, material processing, and specific database entry. In this manner, materials contributing significantly to overall model uncertainty are progressively specified from level 1 (material category) to 5 (specific database entry).

[2] at the use location, not the manufacturing location.

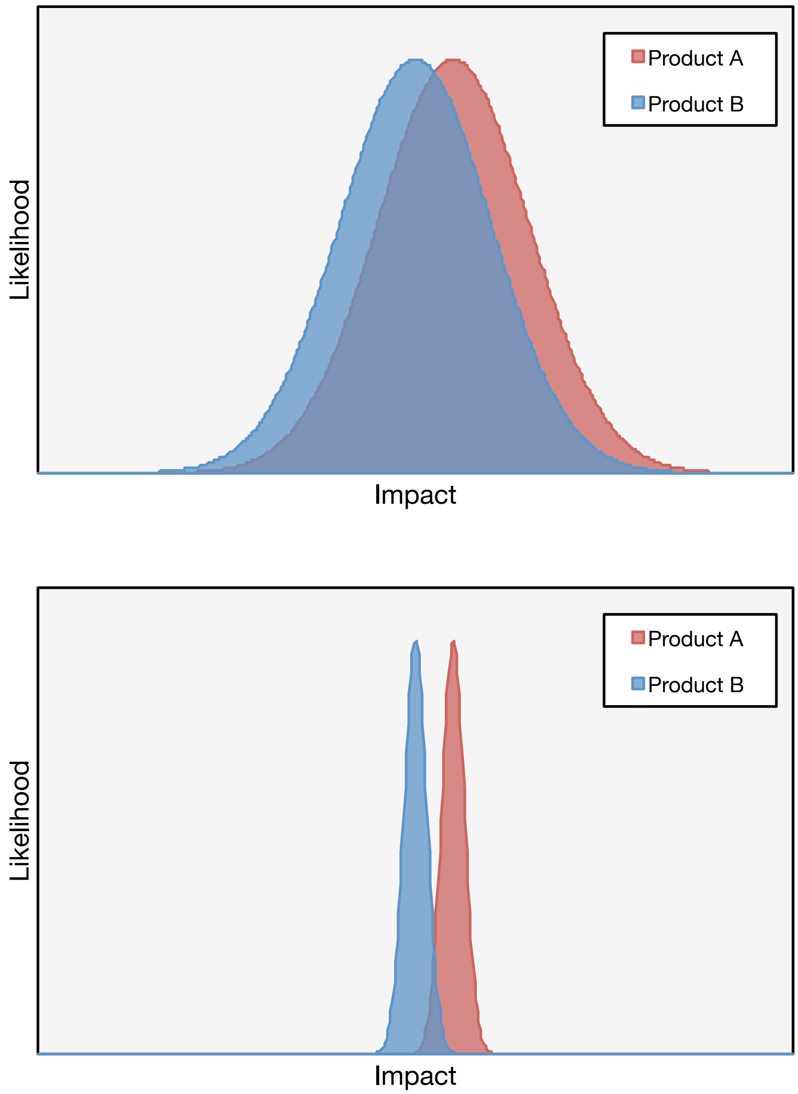


Figure 2.1 – Illustration of the uncertainty reduction process in streamlined life cycle assessment, before and after resolution. This figure depicts the probability distributions of the impact of two different products. Above are the distributions before uncertainty reduction. Below are the distributions of the same products after resolution of parameters; note that the intersection of the two distributions is markedly smaller, and therefore the impacts are more easily differentiated.

2.3.4 System Parameters for Monte Carlo Simulations for LCA

The life cycle assessment case studies mentioned in this section use Monte Carlo simulation for a wide variety of system parameters and their variability or uncertainty. Below is a sample of parameters.

Material mass

The mass of certain parts may vary or be uncertain. Because the embodied energy of materials can in many cases account for a large part of the environmental impact, any uncertainty or variability in the mass must be accounted for in the model.

Material composition

Similarly, the material composition may be somewhat unknown. The range of materials that could be used must be accounted for, as in Patanavanich's [21] material underspecification method.

Use location

The geographic location where a product is used may have a proportionally large environmental impact.[3] Local characteristics, such as the mode of electricity generation, may significantly affect the use phase impact; therefore, any uncertainty about location must be modeled.

Use duty cycle

The duty cycle some products are subjected to may vary tremendously. For example, a laptop may be used by an employee for 8 or more hours a day; the same model, however, may be a second personal computer in a home and used infrequently. As mentioned above, if the use phase has a significant affect on environmental impact the variability of its duty cycle must also be modeled.

[3] Indeed, for most products that actively consume energy during the use phase to provide their primary functionality, the use phase will dominate the life cycle impact, especially with regards to energy [28].

Product attributes

For many products, fundamental attributes may be appreciably variable. For example, a load-bearing part may be optionally composed of either 1 kg steel or 2 kg of aluminum. If these attributes are not known, they may be simulated. One method of determining the distribution for such parameters is the market share. For example, if 75% of the products in question utilize steel, then the probability of sampling steel for this parameter will be 0.75.

If an LCA covers a range of products as opposed to a specific model, even more fundamental product attributes may vary, and therefore should be modeled with uncertainty. An example from one of the case studies is the size of an LCD display.

2.4 Industry Collaboration: Diesel Engine Manufacturing

This research project focuses on the application of life cycle assessment to diesel engine manufacturing. A large international engine manufacturer collaborated on this project, contributed to its direction, and provided a necessary test opportunity for LCA with one of their engine models.

2.4.1 Engine Studied

The engine studied in this research is a large diesel truck engine, with an engine displacement of approximately 15 liters. The primary application of the engine is on-highway tractor-trailers, with gross vehicle weight ratings of 25,000 lb. to 80,000 lb.[4]

The engine is sold both domestically and internationally, and is therefore subject to a variety of international regulations; in the United States, these regulations are the US Environmental Protection Agency's (EPA) National Clean Diesel Campaign [30].

The engine is a complex assembly. The bill of materials (BOM) of the engine consists of more than 500 distinct parts, which are comprised of over fifty materials and

[4] The US Department of Transportation's Federal Highway Administration classification of these trucks is Class 7 and 8, or "Heavy Duty." [29] Examples are tractor-trailers, city buses, and dump trucks.

subjected to a variety of processes around the world, both at supplier locations and the in manufacturer's own facilities.

2.4.2 Diesel Engine Production

This project specifically focuses on the production and manufacturing of diesel engines, as opposed the use of the engines.[5] Currently the world market for newly produced diesel engines is estimated to be roughly 150 billion USD [31]. In the United States, the heavy-duty diesel engine manufacturing sector currently employs over 10,300 individuals in 30 different corporations; the sector has revenues of about 11 billion USD [32]. Sales in the Unites States of large diesel vehicles (FHWA Class 8, "heavy duty" [29]) are rising, roughly doubling to 170,000 from 2009 to 2011; this is even greater than before the 2008-2010 recession (150,000 such trucks were sold in 2007) [33].

[5] In this study the extraction and manufacturing phases of a diesel engine's life will be investigated. The in-use phase of its life, however, has a significant environmental impact; in fact, in many respects its impact dwarfs that of the production phase. This observation and the motivations behind the decision to focus on production are discussed in Section 3.2.3 on page 37.

Chapter 3

Modeling and Simulation Methodology

The methodology for this research project is described below. A model will be built to quantify the environmental impact of the production of a diesel engine. This model will have high uncertainty; a simulation procedure will then be run to identify which parameters should be specified more precisely to reduce the overall uncertainty.

3.1 Overview of Process

The general steps of model creation and simulation are given below, along with examples from this study. Each item will be more fully described in this chapter.

3.1.1 Simplified Process Steps

1. Define the scope of the life cycle analysis and the impact metric or metrics used for evaluation. In this study the scope of the LCA is materials extraction through final assembly of the engine. The primary impact metric is embodied energy.

2. Inspect the product's bill of materials, list all major components, and give overall product specifications.

3. Select a subset of the components that are estimated to have the highest impact. In this study, the subset was chosen by cost.

4. Parameterize the major components of the engine. Approximate product specifications, materials, and processes as best as possible. In this study this was done for all of the major components, as well as the remainder of the engine (*en masse*).

5. Use LCA databases to find the environmental impact of the system as a function of the parameters. Treat the LCA values themselves as parameters of the model. An example of parameters in this study are the mass and material composition of the engine block.

6. Bound uncertainties for all parameters loosely based on readily available data. For example, the mass of the engine block (before any material removal) was loosely bounded from technical drawings of the part.

7. Execute a round of simulations. A round consists of a number of separate simulations. Each simulation is the process of randomly setting each parameter to a value in its range and then calculating the impact of the engine. In our example, the engine block mass would randomly be assigned a value in its bounded range. Each round consisted of 10,000 simulations.

8. Calculate the variance (a proxy for the uncertainty) of the impact values calculated by the simulation.

9. Using the output of the simulations, identify the parameter that most significantly affects the uncertainty of the impact. In this study, the partial rank correlation coefficient was used to judge the relative significance of parameters.

10. Reduce the uncertainty associated with this parameter to a reasonable level; this uncertainty reduction represents the process of collecting more specific data on the parameter. For example, the uncertainty of the engine block's mass could be reduced by referring to records of the casting's measured mass. This specific data would then be used to construct a tighter bound on this parameter.

11. Repeat steps 6 through 9 until uncertainty of the impact has either reached an acceptable state, or has reached a steady state across simulation runs.

3.1.2 Workflow and Tools

The model for this project was assembled in Microsoft Excel. The random simulations were executed using a macro script inspired and derived from Auer [34]. All analyses were completed in the R environment [35], with the uncertainty analysis functions provided by the "Sensitivity" package [36].

3.2 LCA Approach, Scope, and Metrics

Before beginning the modeling and data collection phases of a life cycle assessment, the scope of the study and the metrics used to evaluate it must be established. In this section the overall approach to LCA used is discussed, along with LCA metrics and scope.

3.2.1 LCA Approach: Use of Streamlining

Many of the techniques mentioned in 2.2 on page 24 by Weitz [17] and Hunt [18] are used throughout this case study. It should be noted that each of these streamlining techniques has recognized disadvantages; in fact, the primary purpose of this research and case study is to identify which streamlining approximations affect the outcome most adversely so that they can be refined. Overall, the approach is to initially model the system in a general fashion and then to increase specificity as needed.

3.2.2 Impact Metric: Embodied Energy

To assess the impact of engine production, energy was selected as the primary metric. Energy will then serve as a proxy for overall environmental impact.[1] A single metric was selected rather than multiple metrics to simplify the case study. This is similar to one of the streamlining approaches used by Hunt [18]. This methodology can, of course, work with multiple impact metrics; this is explored more fully in Section 6.3.2 on page 92.

There are a number of reasons for the popularity of energy as an environmental metric. Energy has been one of the primary metrics for many environmental studies. For example, energy is the primary metric in Smith et al. [37], Gutowski et al. [11], and Patanavanich et al. [21] Many other environmental indicators show some correlation with energy usage. In fact, Huijbregts et al. [38] make an argument for the use of cumulative energy demand (CED) from fossil fuels as a proxy environmental indicator:

[1] The usefulness of energy as a proxy in this specific study is analyzed more closely in Section 5.1.2 on page 71.

> "Fossil energy demand is indicative of many environmental problems. Fossil CED can therefore be used as a screening indicator for environmental performance instead of performing a full LCA, for instance, in the absence of sufficient data."[38]

Energy, more than other environmental indicators (such as greenhouse gas emissions or water usage), is most closely associated with production costs. Because of this, an energy assessment may have usefulness for a corporation far outside of environmental responsibility. There may therefore be more motivation and resources to study energy. The validity of embodied energy as a proxy for other indicators in this specific case study is discussed in more detail in the results section of this work, Section 5.1.2 on page 71.

There are various methods to quantify energy use. For this study the concept of embodied energy is used. Embodied energy is the amount of energy used to produce a specific quantity of a material, good, or service [9]. The database used for this study is Ecoinvent [39], discussed more in Section 3.3.1 on the following page. Ecoinvent defines embodied energy as the total energy input into a product (and to any input materials and subassemblies); this energy can be derived from any of the following sources: coal, oil, natural gas, nuclear, hydroelectric energy, wood, wind, photovoltaic energy, solar heat, and biofuels [40]. Unless otherwise specified, "energy use" in this work refers to embodied energy.

3.2.3 LCA Scope: Engine Production

As mentioned earlier, this project is a collaboration with a large international diesel engine manufacturer; one of the manufacturer's own engine models served as the target for this LCA case study.

The specific engine studied is a large diesel truck engine, with an engine displacement of approximately 15 liters. The primary application of the engine is on-highway tractor-trailers, which are classified by the US Department of Transportation's Federal Highway Administration as Class 7 and 8 [29]. The engine assembly has more than 500 distinct parts, which are comprised of over 50 different materials and undergo a large variety of processes.

The use phase of an engine is known to dominate most of its life cycle impacts. Keoleian [41] and Smith [37] have shown that for automobiles the use phase dominates the total energy consumption of the vehicle and engine, accounting for 70-80% of the total. As a diesel truck engine may cover 5 to 10 times the distance as an automobile in its lifetime[2], and the fuel consumption per distance is higher, this proportion will likely be even greater. Rough calculations for this engine indicate that the use phase is most likely 99% or more of the total lifetime embodied energy of the engine. The National Academies estimated the typical fuel economy for diesel trucks of this size in the range of 2.5 to 7.5 miles per gallon [42]. An engine with a lifetime travel of 500,000 miles[2] in a typical truck would consume 66,600 to 200,000 gallons; the embodied energy in the fuel alone would then be 10.2 to 30.6 TJ.[3] As discussed in Section 5.1, the embodied energy for the production of the case study engine is approximately 69.7 GJ, or 0.2% to 0.7% of the fuel's total embodied energy.

This case study specifically focused on the first life cycle stages of the engine: material extraction, material production, manufacturing, and part transportation. It must be noted that the use phase was not included in the life cycle assessment. Because engine efficiency is such a significant factor in their business, our industry partner had existing dedicated projects to study use-phase impacts. This case study was then designed to focus on other aspects of the engine's life cycle, namely every stage prior to use. End-of-life stages, such as remanufacturing, recycling, and disposal were also not considered in this study.

3.3 Data Sources

3.3.1 Ecoinvent Database

The primary database used for impact assessment was the latest release of the Ecoinvent database, version 2.2, which was published in 2010 by the Swiss Centre for Life Cycle Inventories [39]. This database was chosen because of its relatively large size, its comprehensive listings, and the availability of international data. The database contains over 4,000 life cycle indicators (LCIs) for energy sources, material

[2]Keoleian and Smith both use nominal values of approximately 100,000 to 120,000 miles for the lifetime distance of an automobile engine before remanufacturing. As a comparison, the diesel engine in this case study has a warranty for 500,000 miles.

[3]The environmental database Ecoinvent [39] (more in 3.3.1) lists the embodied energy of diesel at regional storage facilities as 54.6 to 55.0 $^{MJ}/_{kg}$. The embodied energy for any fuel will be more than its energy content, as all energy expenditures are involved, such as extraction and transportation. The US Department of Energy estimates the energy content of #2 diesel fuel as 137,000 $^{BTU}/_{gal}$ or 45.6 $^{MJ}/_{kg}$.

extraction, transportation, processing, and disposal. Many indicators are available for multiple geographies, such as the European Union and the United States.

The Ecoinvent database is integrated into the life cycle assessment software package that was evaluated as part of this research project: the Windchill LCA module by PTC. The Windchill LCA software is discussed in detail in the appendices, along with an assessment of its utility and recommendations on usage (see Appendix A on page 105).

3.3.2 Other Sources

Not all materials and processes needed for this case study were found in the Ecoinvent database. To construct estimates for the values and ranges of these parameters, two other sources were consulted. Ashby [9] has collected approximate embodied energy data for many different materials. Hammond et al. [43] have similarly created a resource that details the historical values and ranges of embodied energy for various materials. In some cases multiple sources were combined to create the best estimate for a parameter. Details of the materials used in this study can be found in Table B.10 on page 126.

3.4 Formulating and Populating the Model

One of the primary goals of this research is to streamline the data collection necessary for modeling the engine's production impact. Data for each model parameter was bounded by reasonable upper and lower limits. It was only further resolved to a more specific range after an explicit resolving step. This mimics the act of specific targeted data collection. Below is the process used for building a model of the engine's impact during production.

3.4.1 Model Formulation

Equation 3.1 describes roughly the production energy for the engine.

$$E_{\text{total}} = \sum_{i \in P} \left(\sum_{j \in M} m_{ij} e_{ij} + \sum_{k \in R} m_{ik} e_{ik} + m_i \sum_{l \in T} d_{il} \left(\frac{e}{d} \right)_{il} \right) \tag{3.1}$$

The following are the parameters of (3.1).

Parts, P The set P represents all the parts in the engine assembly. Each individual part is expressed as $i \in P$.

Materials, M The set M represents all of the materials used in the engine assembly. Each unique material is expressed as $j \in M$.

Processes, R The set R represents all of the processes used in the production of a part. Each unique process is expressed as $k \in R$.

Transportation Modes, T The set T represents all the distinct modes of transportation used to transport a part to the final assembly location. Each individual mode is expressed as $l \in T$.

Mass, m Mass is represented by m, and is typically given in kg. The mass of each part $i \in P$ is expressed as m_i.

Energy Intensity, e Energy intensity, or specific energy, is represented as e and is measured in energy per mass, typically $^{\mathrm{MJ}}/_{\mathrm{kg}}$. For each part $i \in P$ and material $j \in M$, the unique specific energy of the material is expressed as e_{ij}. Similarly, the specific energy for part $i \in P$ and process $k \in R$ is expressed as e_{il}.

Travel Distance, d The travel distance of each part to the final assembly location is represented as d and is measured in km. For each part $i \in P$ and mode $l \in T$, the distance traveled is expressed as d_{il}.

Energy Intensity of Transportation, $^{e}/_{d}$ The energy intensity of a mode of transportation is represented as $^{e}/_{d}$ and is measured in energy per mass per distance, $^{\mathrm{MJ}}/_{\mathrm{kg \cdot km}}$. For each part $i \in P$ and mode $l \in T$, the specific energy of transportation is expressed as $(^{e}/_{d})_{il}$.

Equation 3.1 can also be expressed in a compact form,

$$E_{\text{total}} = \sum_{i \in P} \left(E_{\text{materials},i} + E_{\text{processes},i} + E_{\text{transit},i} \right) \tag{3.2}$$

where the total energy E_{total} is the summation of the energy use contributions from materials, processes, and transit associated with each part.

3.4.2 Bill of Materials

The engine's bill of materials (BOM) was obtained from the manufacturer for analysis. The BOM for the full engine assembly contains 452 individual part numbers; each of these parts is featured in the engine one or more times. Some parts, such as small fasteners, are used in quantities over 20. In total, the engine assembly consists of 1,029 separate parts.

3.4.3 Truncation of the Bill of Materials

Conducting an analysis for all 452 parts would be prohibitively time-intensive; therefore, the list of parts to be analyzed was truncated based on cost. The relative costs for all parts were compiled, and the 38 most expensive part numbers were culled for detailed analysis. For parts appearing more than once, the total cost of all instances was used. The cost of a component was calculated as the final cost to the manufacturer. If a component was produced internally this cost consisted of material costs, processing costs, and labor costs. For components purchased externally, this cost was simply the purchase price.

Cost was chosen as the sorting metric as opposed to mass, which has been used by other researchers, such as Weitz[17]. Cost was a preferred selection criteria as cost tends to scale with size, scarcity, and complexity. As an example, included in the final culled analysis list were the engine block, a shutoff valve, and the engine control unit. The engine block is costly due to its large relative size; its mass is approximately 356 kg. The shutoff valve is costly due to the scarcity of its components, namely the gold used in plating. The engine control unit is costly due to its complexity and the scarcity of its constituent materials; circuit boards routinely contain many different materials, such as C-H-O polymers, halogenated polymers, copper, gold, beryllium, mercury, silica, and alumina [44].

The lack of diversity in the excluded, lowest-cost parts is another justification for the usage of cost as the selection criteria. Of the 402 unique parts not included in the analysis, most fell under a small number of categories. Over 70% of the remaining were either fasteners or hose components, 45.5% and 25.1%, respectively. Fasteners included screws, washers, and nuts; hose components included tubes, gaskets, seals, and connectors. To calculate total impact of the entire BOM, the impact of these excluded parts must be approximated.

The 38 parts not excluded will be referred to subsequently as the “specified parts”. Equation 3.2 on page 40 must be modified to reflect the truncation. The set $P' \subset P$ represents all of the specific parts. The remainder of the engine is treated as a single entity, much like a part, and is represented as q. Equation 3.3 below is the modified version of Equation 3.3.

$$E_{\text{total}} = E_{\text{materials},q} + E_{\text{processes},q} + E_{\text{transit},q} + \sum_{i \in P'} \left(E_{\text{materials},i} + E_{\text{processes},i} + E_{\text{transit},i}\right) \tag{3.3}$$

In this work, the set P will be considered as $P = P' \cup q$, as the remainder of the engine, q, can in many ways be treated as a part, although most of its characteristics cannot be resolved.

3.5 Modeling Parameter Uncertainty

Each of the parameters in Equation 3.1 and Equation 3.3 has inherent uncertainty associated with it. Before resolution this uncertainty is significant. Even after resolution this uncertainty is noticeable and must be accounted for in the model.

Monte Carlo simulations use probability distributions to model the uncertainty of a parameter. For example, the uncertainty of a parameter could be modeled as a normal distribution with a defined mean and standard deviation. Upon resolution the mean may move, but more importantly the standard deviation will be reduced. On the scale of the aggregate model, it is this reduction in standard deviation that will create increasingly narrow impact distributions, as illustrated previously in Figure 2.1 on page 30.

Table 3.1 on the next page lists some of the most common probability distributions for Monte Carlo Simulation. The subsequent sections give more detailed descriptions of the distributions and their uses.

3.5.1 Common Distributions

A number of statistical distributions may be used for Monte Carlo simulation of life cycle assessments. Heijungs and Frischknecht [45] provide an introduction to the use

Distribution	Continuous	Non-Negative	Minimum Data Points
Uniform	Yes	Yes	2
Triangular	Yes	Yes	3
Normal	Yes	No	3
Log-Normal	Yes	Yes	3
Beta (PERT)	Yes	Yes	3
Two-Sided Power	Yes	Yes	3

Table 3.1 – Probability distributions for Monte Carlo simulations. Each distribution has unique characteristics which determine its usefulness for certain applications. In this research it was necessary to have a continuous and non-negative distribution with low minimum data requirements.

of statistical probability distributions in the context of life cycle assessment. The paragraphs below contain descriptions of the most common distributions, in rough order of increasing data availability demands. The probability density function $f(x)$, the expected value $E(X)$, and the variance $\mathrm{Var}(X)$ are also given for each.[4]

Uniform Distribution

Uniform distributions are utilized for parameters with few reliable data points. Maximum and minimum values are selected to bound the parameter. The uniform distribution has no distinct mode, as all values in the range are equally possible. The distribution's mean is simply the average value of the two extremes, which may also be undesired. The probability density function for the uniform distribution is

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } x \in [a, b] \\ 0 & \text{otherwise} \end{cases}$$

where a and b are the minimum and maximum values in the range.

The expected value of the distribution is

$$E(X) = \frac{1}{2}(a + b)$$

and variance is

[4] See Morgan and Henrion [46] for more details on these probability density functions and their applications to uncertainty analysis.

$$\mathrm{Var}(X) = \frac{1}{12}(b-a)^2.$$

Triangular Distribution

Triangular distributions have three parameter values: minimum, maximum, and mode. This distribution is very common in life cycle analysis Monte Carlo simulation. The primary advantages are its simplicity, non-negativity, and relative descriptive power (when compared with a uniform distribution). Triangular distributions may also be used instead of a normal or log-normal distribution if the data points have an apparent skew.

The probability density function for the triangular distribution is

$$f(x) = \begin{cases} \frac{2(x-a)}{(b-a)(c-a)} & \text{for } x \in [a, c] \\ \frac{2(b-x)}{(b-a)(b-c)} & \text{for } x \in (c, b] \\ 0 & \text{otherwise} \end{cases}$$

where a, b, and c are the minimum, maximum, and mode of the distribution.

The expected value of the distribution is

$$E(X) = \frac{1}{3}(a+b+c)$$

and variance is

$$\mathrm{Var}(X) = \frac{1}{18}(a^2+b^2+c^2-ab-ac-bc).$$

The triangular distribution is an asymmetric distribution, but can be made symmetric if

$$c = \frac{b-a}{2}.$$

Normal (Gaussian) Distribution

Normal distributions are assigned to parameters with more reliable data points. Specifically, a normal distribution must have two parameters, a mean and a standard deviation. These values must be obtained from a dataset of no less than three values.

Normal distributions are not guaranteed to be non-negative, and therefore may not be applicable for some parameters. When using a normal distribution, the underlying data points should be relatively unskewed.

The probability density function for the normal distribution is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

where μ and σ are the mean and the standard deviation of the distribution.

The expected value of the distribution is simply

$$E(X) = \mu$$

and variance is

$$\mathrm{Var}(X) = \sigma^2.$$

Log-Normal Distribution

Log-normal distributions are quite common in life cycle assessment simulations. This distribution is also ascribed to data points in many life cycle indicator databases, such as Ecoinvent. The log-normal distribution uses two parameters: mean and standard deviation. The primary advantage of a log-normal distribution over a normal distribution is the log-normal's non-negativity.

The probability density function for the log-normal distribution is

$$f(x) = \begin{cases} \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(\ln x-\mu)^2} & \text{for } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

where μ and σ are the mean and the standard deviation of the variable's natural logarithm, $\ln x$.

The expected value of the distribution is simply

$$E(X) = e^{\mu + \frac{\sigma^2}{2}}$$

and variance is

$$\mathrm{Var}(X) = (e^{\sigma^2} - 1)e^{2\mu+\sigma^2}.$$

3.5.2 Less Common Distributions

The following two distributions are less common in Monte Carlo simulations for LCA.

Beta Distribution

The beta distribution is occasionally used in life cycle assessment simulations, generally as the specific PERT distribution. The PERT distribution was originally developed for PERT analysis: Program Evaluation and Review Technique. Like the uniform and triangular distributions, it is bounded and non-negative.

Two-Sided Power (TSP) Distribution

The two-sided power (TSP) distribution was introduced in 2002 by J. René van Dorp and Samuel Kotz [47]. It is based on the triangular distribution and is designed to replace distributions such as the PERT. Like the triangular distribution, the TSP distribution is bounded and non-negative.

3.5.3 Distribution Selection

As discussed previously, each parameter in the model of the engine impact was assigned a range. This range was then used to construct a simple uniform distribution for each parameter. The uniform distribution was selected for its simplicity and wide applicability. Uniform distributions have been used in similar Monte Carlo simulations [23, 48], and reflect a high degree of parameter uncertainty, especially when a more specific distribution is not known.[5] In this project, many parameters were estimated from engineering knowledge and therefore were not described by a probability distribution; an example is the mass of material removed from a part by a machining process. Other parameters, such as transportation distance for individual

[5]Patanavanich [21, 22] uses a similar but distinct approach to model high uncertainty and lack of knowledge about a probability distribution.

parts, could be characterized in a Bayesian manner by analyzing all similar instances on record; however, this would require additional data gathering that may be out of the scope of an impact analysis.

3.6 Model Parameters

The final engine model consisted of 184 parameters. Each parameter was assigned an associated uncertainty range, from which data points were sampled in the Monte Carlo simulations, covered in a subsequent section. The uncertainty range of each parameter was intended to reflect the actual initial uncertainty associated with a realistic life cycle assessment or impact analysis. Each of the ranges was intentionally wide, and was derived from existing sources, such as the Ecoinvent database, the engine's bill of materials (BOM), and technical drawings of the constituent parts.

3.6.1 Material energy intensity

The engine BOM contains material information for the constituent parts. This material information is varied, from specific ANSI material declarations to general ranges for acceptable materials. To analyze the impact of extracting and preparing these materials, representative materials were chosen. In many cases, a single representative material could be selected from the Ecoinvent database. In cases in which a single material was not available, substitutes were made. These substitutes relied on other reputable sources. In every case, an effort was made to quantify the range of uncertainty in the material data for each part. These uncertainties come from the variability in materials used for a part, the variability in the Ecoinvent material data, and the variability in material extraction and processing locations.

Forty-three parameters described the material intensity of the parts and the remaining composition. For each part $i \in P$ and material $j \in M$, the unique energy intensity of the material is e_{ij}. The ranges for most of the materials was obtained from the Ecoinvent database [39]. For materials not covered by Ecoinvent, data from Ashby [9] and Hammond et al.[43] was used to create a feasible range of values. The unresolved range of values for each material spanned the reasonable range for the lowest to the highest values available. Table B.10 in the Appendix details the range of material intensity as well as the data source for each material.

3.6.2 Part mass

43 parameters described the mass of specified parts and the remainder of the engine. The range of the mass for each specified part was derived from technical drawings and associated information about the material composition. This information was in reality much easier to obtain than specific part mass, and therefore is a realistic reflection of the streamlining process. The total engine mass was bracketed by the masses of similar engines in the same classification. The mass of each part $i \in P$ is expressed as m_i.

3.6.3 Process masses

The material processing data available was commonly more general than the material extraction data. For example, the Ecoinvent database contains nine different indicators for the extraction of various steel alloys; however, the same database lists only one general heat treatment process for steel. Processing of the engine components were divided into various categories:

Material shaping Material shaping processes do not change the mass of the part. These processes are sometimes referred to as "net shape." Material shaping processes include forging and bending.

Material removal Material removal processes reduce the mass of the part through the processing step. Examples of such processes are milling, grinding, and drilling.

Material addition Material addition processes increase the mass of the part. Similar or different materials may be added. Examples of these processes include plating, coating, and painting.

Twenty-four parameters specified the mass acted on by various processing steps. The range for these masses was derived from technical drawings of the parts, in the same manner as the part mass estimations. It should be noted that additional processes were applied to the entire part, and therefore were directly tied to part mass.

3.6.4 Remaining material composition

Five parameters characterized the composition of the parts not specified. The composition of the remainder was divided into six general materials:

- Low-alloy steel
- Stainless steel
- Cast iron
- Tin
- Titanium
- Cast aluminum alloy

The five parameters described the composition of each of these materials, with the exception of low-alloy steel, which was the dominant material and calculated as the remainder.

This information was calculated by the manufacturer from purchasing records. The manufacturer uses this aggregate information to monitor the cost of the material input to the engine. This information was very beneficial when assessing the impact of the engine; after the most expensive parts were defined, the remainder of the engine can be treated as a aggregate part, with known material composition, and uncertain processing.

3.6.5 Process intensity

Forty-one parameters described the energy intensity of manufacturing processes acting on individual parts. The range of energy intensity values for these processes was gathered from the Ecoinvent database. The energy intensity of processes is usually given in $^{MJ}/_{kg}$, energy expended per mass processed.

3.6.6 Part transportation distance

Four parameters characterized the distance individual parts traveled to the assembly location after fabrication. The three heaviest parts were treated individually: the

Mode	Lower e/d, $\mathrm{MJ}/\mathrm{t{\cdot}km}$	Upper e/d, $\mathrm{MJ}/\mathrm{t{\cdot}km}$
Truck	1.7	7.3
Rail	0.49	0.70
Ship	0.15	0.61

Table 3.2 – Range of transportation mode energy intensities. Note that these values are given per metric ton, rather than kg. More details are given in the Appendix in Table B.2 on page 113.

cylinder block, the cylinder head, and the engine crankshaft. These three parts alone account for 756 kg, or 54%, of the total engine mass. The remained of the parts were treated as a single mass.

The ranges of transportation distance were estimated from the manufacturer's supply chain. Some parts were processed internationally in multiple continents, while others were sourced domestically.

3.6.7 Transportation mode and intensity

Twenty-four parameters described the transportation mode and intensity. Three transportation modes were available: on-road trucks, rail transit, and ocean and sea vessels. The mix of these modes was also parameterized and allowed to vary. The energy intensity associated with each mode was bracketed by ranges from the Ecoinvent database. This intensity is represented as e/d and measured in energy per mass per distance, $\mathrm{MJ}/\mathrm{t{\cdot}km}$. Table 3.2 lists the three transportation modes and their associated ranges.

3.7 Assessing Model Uncertainty and Selecting Parameters for Resolution

After the model was constructed and all parameters were assigned ranges, Monte Carlo simulations were carried out to characterize the uncertainty at each model state. Model parameters were progressively resolved and made more specific. After a parameter was resolved, the model was simulated again. Each round of model simulation consisted of 10,000 separate model simulations, in which each of the 184 model parameters was selected from its random distribution.

3.7.1 Model Uncertainty Metrics: False Signal Rate from the Self-Test

The uncertainty associated with each model simulation was characterized using a "self-test" inspired by Zgola [23]. The primary aim of the self-test is to characterize the amount of resolution available in the current state of the model. In the self-test the distribution of impact values from a simulation round are duplicated and scaled. Let $X_A \in \mathcal{X}$ be a random variable that represents the observed impact distribution, and $\widetilde{X}_A \subset \mathbb{R}_+$ be a set of draws from X_A. Let $\widetilde{X}_B \subset \mathbb{R}_+$ be the set of the duplicated self-test, defined as

$$\widetilde{X}_B = \alpha \widetilde{X}_A$$

where $\alpha \geq 1$ is a constant scalar and therefore $E(\widetilde{X}_B) \geq E(\widetilde{X}_A)$. The random variable X_B is then never drawn from, but rather $\widetilde{X}_B$ is calculated explicitly,

$$\widetilde{X}_B = \{x_B | x_B = \alpha x_A \ \forall x_A \in \widetilde{X}_A\}$$

The false-signal rate (FSR) is calculated using $\widetilde{X}_A$ and $\widetilde{X}_B$, and is a measure of the confusion between the two:

$$FSR(\widetilde{X}_A, \widetilde{X}_B) = P(\widetilde{X}_B < \widetilde{X}_A)$$

Again, note that $E(\widetilde{X}_B) \geq E(\widetilde{X}_A)$. Because these are both discrete simulations, FSR can be calculated explicitly,

$$\begin{aligned}
FSR(\widetilde{X}_A, \widetilde{X}_B, \alpha) &= \frac{1}{|\widetilde{X}_A||\widetilde{X}_B|} \sum_{x_A \in \widetilde{X}_A} \left(\sum_{\{x_B \in \widetilde{X}_B | x_B < x_A\}} 1 \right) \\
&= \frac{1}{|\widetilde{X}_A|^2} \sum_{x_A \in \widetilde{X}_A} \left(\sum_{\{x_B \in \widetilde{X}_B | x_B < x_A\}} 1 \right) \\
FSR(\widetilde{X}_A, \alpha) &= \frac{1}{|\widetilde{X}_A|^2} \sum_{x_A \in \widetilde{X}_A} \left(\sum_{\{x_B \in \alpha\widetilde{X}_A | x_B < x_A\}} 1 \right) \qquad (3.4)
\end{aligned}$$

In simple terms, FSR is the count of all $x_A \in \widetilde{X}_A$ and $x_B \in \widetilde{X}_B$ where $x_B < x_A$, normalized by the total number of comparisons, $|\widetilde{X}_A||\widetilde{X}_B| = |\widetilde{X}_A|^2$. Note that in the final formulation of FSR in Equation 3.4, only $\widetilde{X}_A$ and α are needed.

For this project, the scaling factor was $\alpha = 1.1$ - an increase of 10%. Therefore, the self-test simulated a scenario in which an engine with a 10% larger impact (B) is compared with the original parameterized engine (A). The FSR is a measure of the confusion between the two, the probability of incorrectly assigning a larger impact to A than to B.

Working from the basis created by Cook [49, 50] for exact calculation of inequalities, Zgola calculates the analytical approximation for FSR. This approximation assumes a Gaussian distribution of impact values; while the sampled impact values do not form a precise normal distribution, when large sample sizes are used they approximate a normal distribution.[6]

$$FSR_{\text{approx}}(k, COV_A) = \frac{1}{2}\left(1 + erf\left(\frac{1-k}{2COV_A}\right)\right) \tag{3.5}$$

Where k is the ratio of expected values of $\widetilde{X}_B$ and $\widetilde{X}_A$,

$$k = \frac{E(\widetilde{X}_B)}{E(\widetilde{X}_A)}$$

and COV_A is the coefficient of variation of $\widetilde{X}_A$,

$$COV_A = \frac{\sigma(\widetilde{X}_A)}{E(\widetilde{X}_A)}$$

The formulation of FSR in Equation 3.5 is not necessary for computation (as Equation 3.4 can be used directly) but it is useful for analysis. Using 3.5 it can be shown that the maximum FSR is 50%, and this occurs when $\alpha = 1$ (i.e. when a distribution is compared to itself).

[6] Figure B.3 on page 125 in the Appendix demonstrates the viability of the normality assumption with a q-q plot. See Wilk [51] for an introduction to q-q plots.

$$
\begin{aligned}
FSR_{\text{approx}}(k, COV_A) &= \frac{1}{2}\left(1 + erf\left(\frac{1-k}{2COV_A}\right)\right) \\
&= \frac{1}{2}\left(1 + erf\left(0\right)\right) \\
&= \frac{1}{2}
\end{aligned}
$$

3.7.2 Distribution of Calculated Metrics: The Bootstrap Method

As described above, the false signal rate can be calculated for each simulation round sample, $FSR(\widetilde{X}, \alpha)$ (see Equation 3.4). In many situations it is helpful to calculate the distribution of this statistic like FSR rather than a singular representative value.

The bootstrap method is a resampling method first proposed by Efron [52] which calculates distribution of a statistic $S(Z)$ from a distribution Z. The distribution of the statistic $S(Z)$ should not to be confused with the distribution it is calculated from, Z. Note that Z is itself typically a sample from a larger population; in Monte Carlo simulations, each simulation round produces the sample set $\widetilde{X}$ from the model X.

The bootstrap calculates the sampling distribution of a statistic $S(Z)$ by resampling with replacement from Z. Let Z_i be an instance of the resample. Each resample is of size n; in total k resamples are taken. The statistic S is calculated for each sample. Let $\mathbf{S}^*$ be the set of all statistics calculated on the resamples:

$$
\mathbf{S}^*(Z) = \{S(Z_i) \mid i = 1, \ldots, k\}
$$

It is critical for the bootstrap method to be effective that the sampling from Z be done with replacement. In this manner, the statistic $S(Z_i)$ for large sample sizes n ($n \approxeq |Z|$) will not simply converge to the singular value $S(Z)$ but will instead provide meaningful insight into the possible distribution of $S(Z)$. Sampling with replacement assumes that all $z \in Z$ are independent and identically distributed (i.i.d.) random variables.

In the same way $S^*(Z)$ is defined let $FSR^*(\widetilde{X}_A, \alpha)$ be the bootstrapped distribution of $FSR(\widetilde{X}_A, \alpha)$:

$$FSR^{*}(\widetilde{X}, \alpha) = \{FSR(\widetilde{X}_i, \alpha) \mid i = 1, \ldots, k\} \tag{3.6}$$

where each $\widetilde{X}_i,\ i = 1, \ldots, k$ is a random sample from $\widetilde{X}$ of size n.

3.7.3 Selection of Parameters for Resolution using Partial Rank Correlation Coefficients

The goal of any streamlining procedure is to reduce the amount of data needed to complete an analysis, so the selection criteria for parameters to resolve is extremely important. Each simulation round begins with the resolution, or specification, of a single parameter. The goal is to select the parameter that will reduce the most uncertainty. Partial rank correlation coefficients[7] (PRCCs) were chosen as an estimation of the relative contribution to uncertainty by each parameter.

The PRCC method was selected because it is applicable to nonlinear system, as opposed to competing methods, such as partial correlation coefficients. The PRCC method was proposed in 1942 by M. G. Kendall [53]. This method applies the concept of the partial correlation coefficients to rank analysis. Partial correlation analysis aims to determine if the observed correlation between two real variables, $x_{1,}x_2 \in \mathbb{R}$ is primarily due not to each other but rather a third variable, $x_3 \in \mathbb{R}$. It is of note that any relationship between the variables in question must be assumed to be linear. Kendall then applied this procedure, with modifications, to ranked values $y_{1,}y_2, y_3 \in \mathbb{Z}_+$, producing the partial rank correlation coefficient. The PRCC method may be extended then to any real value variables $z_{1,}z_2, z_3 \in \mathbb{R}$ assuming they have monotonically increasing or decreasing relationships (rather than a more restrictive linear relationship) and can be ranked by value: $\hat{z}_i = \text{rank}(z_i),\ i = \{1, 2, 3\}$.

The PRCC method became a commonplace tool for uncertainty and sensitivity analysis, with the primary benefit being that it does not require a probability distribution to be assumed as only ranks are used. For illustrations of this application, see the work of Ronald Iman and Jon Helton [54, 55]. Some uncertainty analysis software packages have used PRCCs to estimate contribution to variance. As an example, Oracle's Crystal Ball software estimates "Contribution to Variance" with

[7] An example of rank correlations used in the life cycle assessment context to find significant parameters is Maurice et al. [7]

the normalized square of the PRCC of each parameter [56]. PRCC has also been used by researchers in the life cycle analysis community; for example, see Zgola [23], who uses Crystal Ball's "Contribution to Variance" to estimate each LCA parameters' contribution to overall model uncertainty.

In this research, the R package "sensitivity" [36] was used to calculate partial rank correlation coefficients for all parameters. The actual change in the uncertainty of the model was calculated at the conclusion of every simulation round.

3.7.4 Accessible and Inaccessible Information for Resolving

One of the primary goals of this research project is to design a data collection strategy for a realistic scenario. One of the primary differences between idealized impact assessments and their real-world instantiations is limited information. For this study we collaborated with the engine manufacturer. This project and the simulations reflect which information the manufacturer had available.

Information that was available to the manufacturer was considered capable of being resolved. That is, with reasonable resource expenditure, the information could be specified. Information and specifications were considered unresolvable if the task was both extremely laborious and poorly specified. Examples of information that could be resolved are the specific mass of individual parts, the mass of the entire engine, the energy intensity of specific processing steps, the mode of transportation for specific parts, and the distance traveled by parts. Examples of information that could not be resolved are the mass of all remaining (unspecified) parts, the distance these parts traveled, or any other detailed information about the unknown parts.

Very little information was considered specified without a resolving step. Even the mass of each part necessitated a resolving step. The logic for this is that the engine's bill of materials is dynamic, and many parts change supplier and specifications throughout the life of the engine design. Some parts may be sourced from multiple suppliers. Some part designs may be the manufacturer's, while others may be the supplier's proprietary design. Given these realities, it became apparent early in the project that even relatively simple specifications such as part mass should not be considered given without effort.

3.7.5 Resolution of Uncertainty

For this study, the process of parameter resolution consisted of reducing the variance of the parameter. Because all parameters were modeled as uniformly distributed random variables, variance reduction was achieved by tightening the bounds of this distribution.

Resolved bounds in this case study still remained somewhat wide, with typical values in the range of 1%-3% of the parameter value. This was done intentionally, so as to not overstate confidence in any resolved values. It should be noted that these ranges are much wider than those of resolved parameters in other studies. For example, in Zgola's case study [23] on LCA uncertainty reduction for liquid crystal displays, many resolved value ranges are less than 0.01% of the parameter values; the narrowest resolved range is 0.00002%. These narrow ranges will report higher overall model resolution than is possible in reality.

Chapter 4

Simulation Results and Analysis

This chapter discusses the results of the primary focus of this work, the iterative reduction of LCA uncertainty through simulations. The results of the LCA itself are discussed in Section 5.1 on page 70.

4.1 Simulation Results

4.1.1 Initial Results and Lessons

After every round of simulation, each parameter's contribution to the model's overall uncertainty was estimated with the partial rank method. Parameters with high relative $|PRCC|$ values are estimated to have the most significant impact on uncertainty, and are therefore candidates for resolution.

Table 4.1 illustrates the PRCC results after the initial simulation, before any resolution of parameters. Of particular note is the relatively small value for total engine mass, ranked 25^{th}. The results of the first 19 rounds of simulation can be seen in Figure 4.1. It is readily apparent that total engine mass is by far the most influential parameter.[1] In light of this, the simulations were restarted, using the basic heuristic gleaned from the previous simulations: total engine mass was resolved before all other parameters. Besides this, the resolutions followed the suggestions of the PRCC results.

[1] Based on the results of the simulation for the total engine mass parameter, it is possible that the model lacks complete monotonicity between the parameter inputs and the output.

$\lvert PRCC\rvert$ Rank	Category	Description	Resolvable
1	Material Composition	Remaining aluminum fraction	Yes
2	Process Mass	Machining of remaining steel	-
3	Material Intensity	Remaining low alloy steel	-
4	Process Intensity	Machining of cylinder block	Yes
6	Transit Distance	Remaining mass	-
13	Transit Mode	Remaining transportation, truck	-
14	Part Mass	Electronic control module	Yes
15	Process Intensity	Machining of crankshaft	Yes
16	Material Intensity	Crankshaft, low alloy steel	Yes
23	Transit Intensity	Remaining transportation, truck	-
25	Part Mass	Total engine mass	Yes

Table 4.1 – $\lvert PRCC\rvert$ values for selected parameters after initial round of simulations. Detailed results can be found in Table B.3 on page 114.

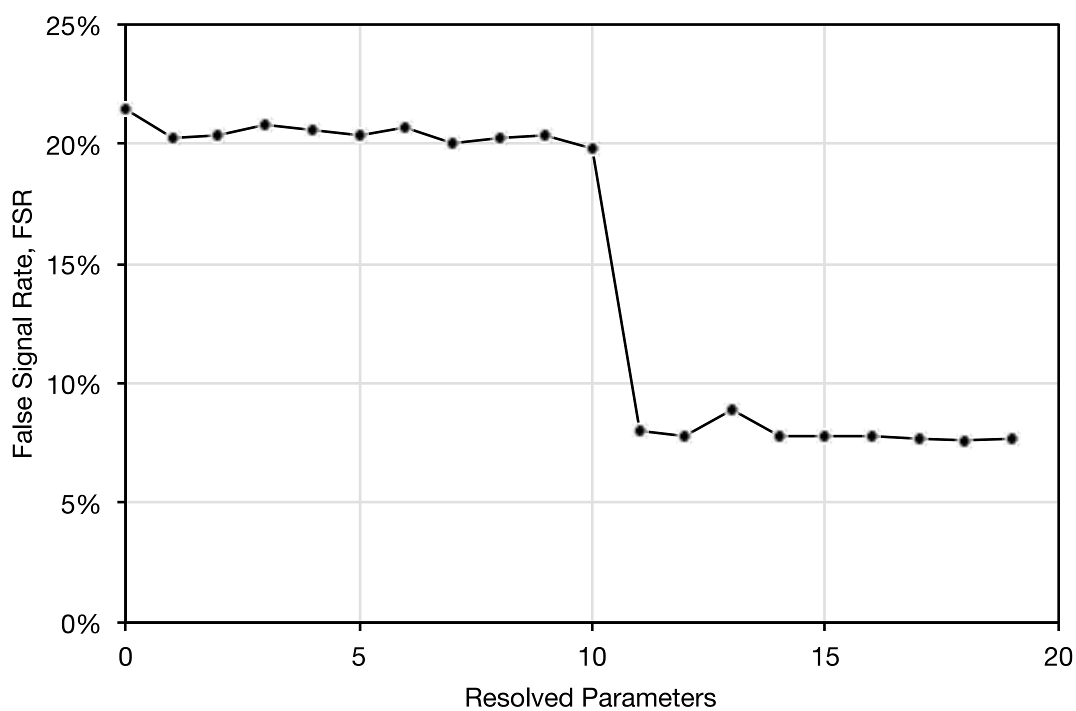


Figure 4.1 – Results of initial first 19 resolved parameters. Note the large decrease in false signal rate after the 11$^{\text{th}}$ parameter, total engine mass, is resolved. Following these results, the simulation was restarted, and total engine mass was resolved first. Those results can be seen in Figure 4.2 and Figure 4.3.

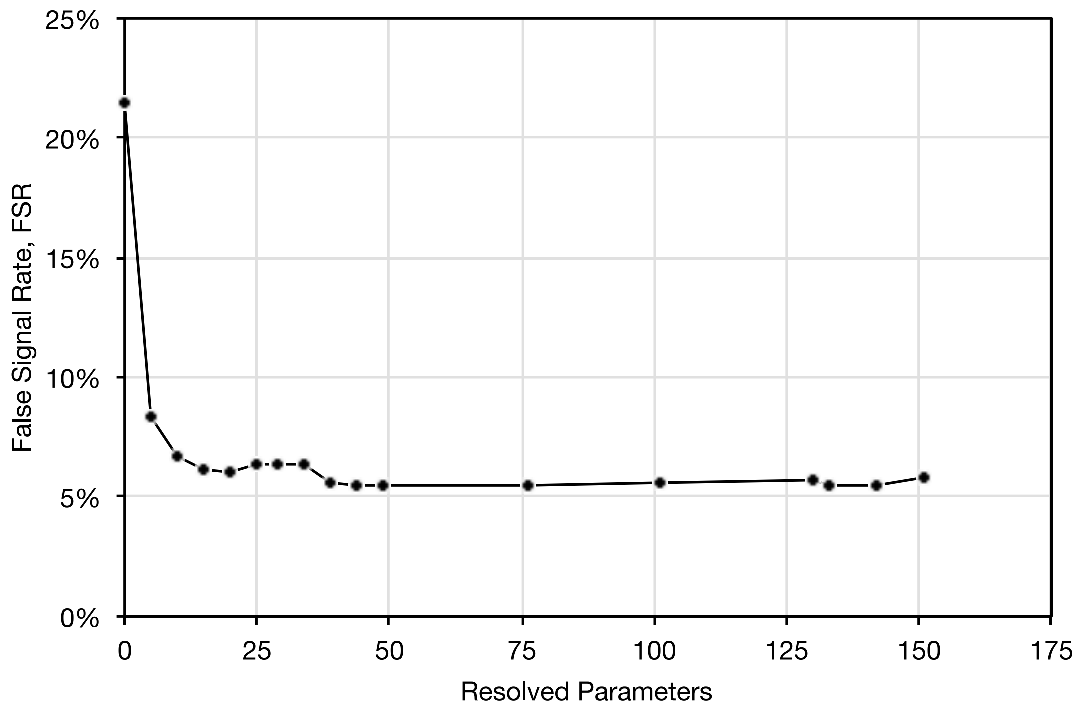


Figure 4.2 – Results of final simulations, all rounds. False Signal Rate is taken from the 10% self-test.

4.1.2 Final Results

Figure 4.2 shows the results of all rounds of final simulations, with Figure 4.3 detailing the first 30 rounds; both figures represent the model after restarting. The model uncertainty is reduced noticeably in the first 13 rounds of simulations. After 39 rounds, however, the model uncertainty is roughly constant regardless of the number of further parameters resolved. Therefore the minimum degree of uncertainty (as quantified by the false signal rate of the 10% self-test) appears to be 5.8% in this formulation of the model. With all 151 possible parameters resolved, the simulated false signal rate was still 5.8%. Figure 4.2 shows the FSR after the resolution of the first 16 parameters, as well as after the resolution of all 151 resolvable parameters.

There is very obvious variation in the results, as illustrated by the apparent increase in uncertainty after some resolution steps. This is partially an artifact of the simulation process. When these results are resampled using the bootstrap technique, it can be shown that each increase in FSR after resolution falls well within the sampling range of the previous iteration's FSR. Figures B.1 and B.2 in the Appendix illustrate this. Also included in the Appendix is Table B.7, which provides various percentiles for

Parameters Resolved	Parameter Category	Parameter Description	FSR
0	-	Initial Round	21.5%
1	Part Mass	Total engine mass	11.9%
2	Material Composition	Remaining aluminum fraction	9.7%
3	Material Composition	Remaining gray iron fraction	9.6%
4	Process Mass	Machining of cylinder block	9.4%
5	Material Intensity	Cylinder block, gray iron	8.4%
6	Material Composition	Remaining titanium fraction	8.4%
7	Process Intensity	Forging of engine crankshaft	7.7%
8	Process Mass	Machining of cylinder head	7.6%
9	Part Mass	Electronic control module	7.3%
10	Process Mass	Machining of crankshaft	6.7%
11	Material Intensity	Crankshaft, low alloy steel	6.4%
12	Material Intensity	Cylinder head, gray iron	6.1%
13	Process Intensity	Machining of cylinder block	6.1%
14	Material Composition	Remaining tin fraction	6.1%
15	Material Intensity	Electronic control module, PCB	6.1%
16	Process Intensity	Machining of cylinder head	6.0%
151	-	All parameters resolved	5.8%

Table 4.2 – False Signal Rate after the first 16 round of simulations. These values were taken after restarting the simulations. Final FSR is also included. Note that the reduction in uncertainty is negligible for the resolution of the last 135 parameters.

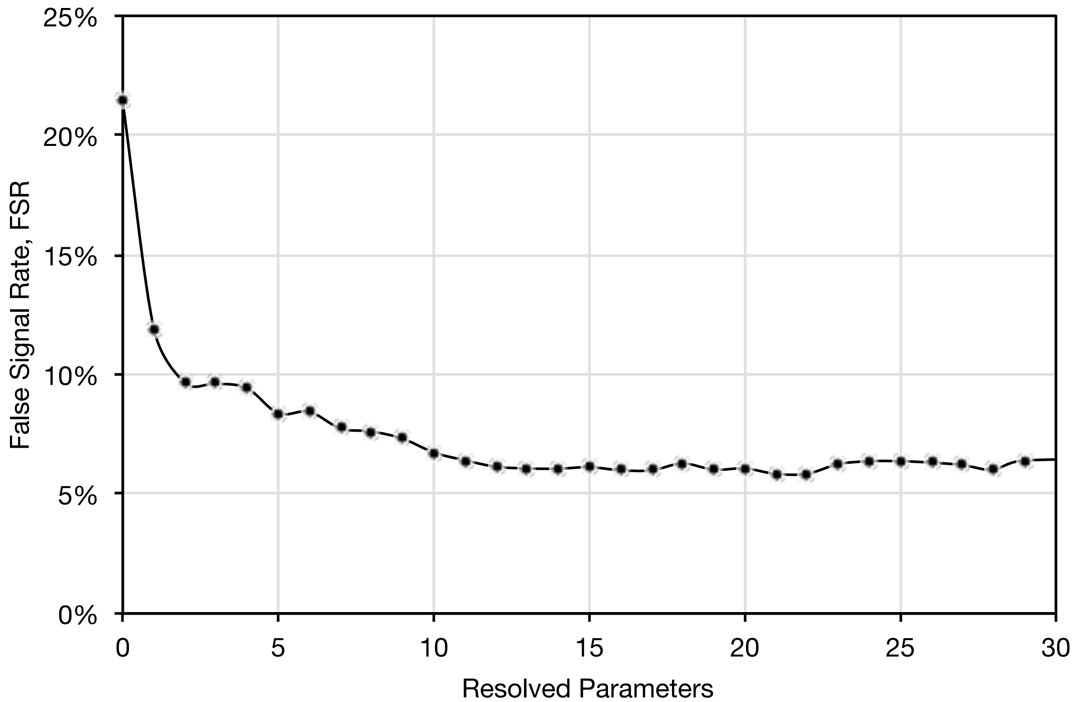


Figure 4.3 – Results of final simulations, first 30 rounds. False Signal Rate is taken from the 10% self-test.

the FSR data. From this table it can be shown that no round's median FSR is more than the 70^{th} percentile of the previous round's FSR. Indeed, only 3 of the 40 rounds have an FSR greater than the 60^{th} percentile of the previous round's FSR.

This phenomenon of increased uncertainty may also be caused by the interdependence of parameters. In a simple example an impact i may be caused by two parts, a and b:

$$i = e_a m_a + e_b m_b \tag{4.1}$$

where i is impact, and e is energy intensity. Expressed in random variables, this is

$$I = E_a M_a + E_b M_b$$

If the sum of m_a and m_b is known (with or without uncertainty), then the random variables M_a and M_b are not independent. Therefore the variance of the impact random variable I may increase upon a decrease in the variance of M_a and M_b, if it is accompanied by a shift in the expected values of m_a and m_b; this exact scenario is possible in a resolution step. This reasoning is explored further in the Appendix on page 128 in Equations B.1 through B.12.

4.2 Discussion of Results

Very few of the parameters in the engine model were found to impact model uncertainty significantly; therefore, these parameters should be targeted for data collection and resolution. This ability to affect the model's uncertainty will be referred to as "leverage" through this thesis.

4.2.1 High Leverage Parameters

Parameters with high leverage included overall approximate engine mass, material composition, part masses, energy intensity of materials, and processing for certain large and expensive parts. A few smaller and less expensive parts affected the model, but typically only via their material energy intensity. Table 4.3 on page 63 shows the 17 parameters that noticeably affected the uncertainty of the model.

The leverage results shown in Table 4.3 are discussed in more detail below.

Part mass

Five mass parameters were observed to have high leverage. Most of these are high-mass parts, above 100 kg: the cylinder head, cylinder block, and crankshaft. One exception was the electronic control module, which was made of a high-energy-intensity material, printed circuit board (PCB). The other exception was the air intake manifold, which was made of cast aluminum; aluminum is significantly more energy intensive per unit mass than iron and steel, the materials of the heaviest components. Also, decreasing uncertainty about the mass of the manifold reduces uncertainty about the mass of aluminum in the remainder of the engine; aluminum was the second most abundant material in the remainder of the engine.

Process intensity

Three of the process intensity parameters with high leverage featured the crankshaft. The crankshaft is unique among the other parts because it has both a large mass and

Category	Part	Details
Part Mass	Total engine assembly	
	Cylinder head	
	Electronic control module	
	Cylinder block	
	Air intake manifold	
Process Intensity	Crankshaft	Forging
	Crankshaft	Machining
	Oil filter head	Machining
	Crankshaft	Heat treatment
	Cylinder head	Machining
Process Mass	Crankshaft	Material removed
	Cylinder block	Material removed
	Cylinder head	Material removed
Material Intensity	Cylinder block	Gray iron
	Crankshaft	Low alloy steel
	Cylinder head	Gray iron
Material Composition	Remaining assembly	Aluminum fraction

Table 4.3 – High leverage parameters, based on decrease in overall model uncertainty, FSR. Leverage is the significance of a parameter's effect on uncertainty and was judged by the median change in FSR from the previous round, as measured in the bootstrap resampling. The above parameters showed a $|\Delta\text{FSR}| > 0.10\%$. For a full listing of the significance levels of all parameters, see Table B.5 on page 117.

many processing steps. The other two parameters in this category also belonged to relatively massive parts (the oil filter head and the cylinder head).

Process mass

All three of the process mass parameters with leverage involved processes that effected large masses. All three processes also involved material removal, which can be very energy intensive.[2]

Material intensity

The three material intensity parameters follow a similar theme: all involve very massive parts. The material itself, however, is not necessarily energy intensive; indeed, the materials involved in these process - gray iron and low alloy steel - were two of the least energy-intensive materials considered in this project.

Material composition

In this model there were five material composition parameters that effected the remaining mass of the engine, and therefore had the potential for high leverage. Only one such parameter was found to have high leverage. It should be noted, however, that resolving this parameter accounted for more than 90% of the remaining engine mass.[3] Therefore, the other material composition parameters did not have any significant leverage.

Themes in high-leverage parameters

A parameter may have large leverage for various reasons.[4]

- Magnitude of uncertainty. If the uncertainty of a parameter is high enough, it may have significant leverage in the model. Example: the mass of a complex part like the air intake manifold was difficult to estimate and therefore had high uncertainty.

[2] In this study material removal processes included the embodied energy of the material removed, as well as the energy involved in the actually removal process.

[3] Low alloy steel and aluminum made up the vast majority of the remaining engine mass.

[4] Geisler et al. [57] discuss in more detail these various factors in a parameter's contribution to uncertainty.

- Magnitude of impact. If the environmental impact of a parameter is high, it may affect the model significantly, even if its uncertainty is relatively low. Example: the cylinder block's mass is very high relative to the engine as a whole, and contributes noticeably to the impact.
- Reuse in model. If a parameter is referenced in multiple sections of the model, it may have high leverage. Example: total engine mass is referenced numerous times in the model. The uncertainty was relatively well bounded, but the reuse and magnitude of this parameter give it the highest leverage.

Relationship to PRCC

It is beneficial to compare parameters recognized as high-leverage with those receiving high PRCC values. This is one method of assessing the usefulness of PRCCs as a predictor of parameter importance in the model. Figure 4.4 on the following page compares a parameter's realized leverage - measured as the decrease in uncertainty, FSR - with its PRCC value. This comparison is revealing; some parameters with low PRCC values had noticeable leverage. Parameters with high PRCC values do, however, appear to be more likely to have leverage. Overall, the relationship appears to be relatively weak.[5]

4.2.2 Insignificant Parameters

As expected, the parameters associated with many of the smaller parts were found to be insignificant. Less intuitive was the finding that the transportation parameters for individual parts were not selected for resolution (i.e. low $|PRCC|$) and did not contribute to model uncertainty in a discernible manner. This is a significant insight: total transportation accounts for a significant portion - roughly 9% - of the total energy used in this study's boundaries, yet its constituent parameters have no appreciable leverage in the model.

In this light, it may be possible to leave all transportation unresolved and still produce LCA results with acceptable uncertainty. It may also be possible to collect aggregate statistics on the transportation of parts to slightly tighten the bounds of all transportation parameters.

[5]It should be noted that only parameters that were able to be resolved (see Section 3.7.4 on page 55) can be compared in this manner. Therefore, many parameters with high PRCC values are left out of this comparison.

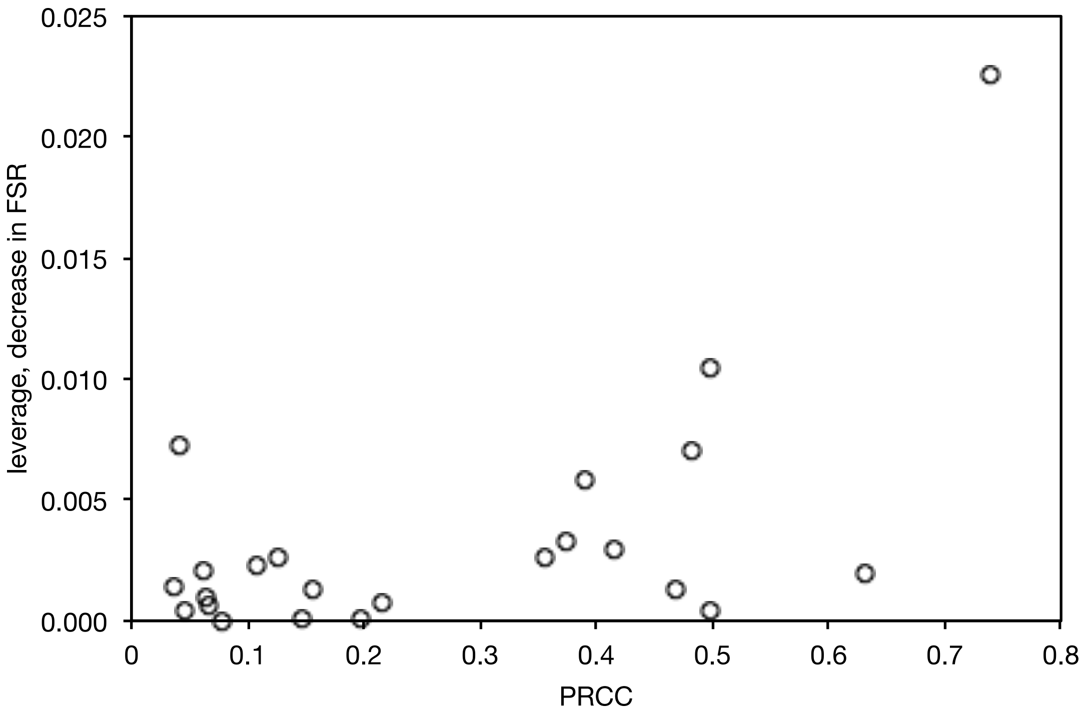


Figure 4.4 – Parameter leverage realized vs. PRCC. Parameter leverage is measured as the decrease in the false signal rate, FSR.

Many materials in the total engine composition were found to be insignificant. Low alloy steel and aluminum made a up a large degree of the remaining engine mass and were the only parameters that showed high leverage. The other materials (e.g. iron, stainless steel, titanium, and tin) could be left unresolved without significant increase in uncertainty.

4.2.3 Inherent Uncertainty

It was shown that the model and the system it represents have significant inherent and unavoidable uncertainty. This point, while obvious qualitatively, has significant ramifications for data collection. The inherent uncertainty in the formulation dominated most of the parameters in the model, and therefore approximately only 40 parameters were significant enough to affect overall uncertainty. After these data points are collected, analysis should either cease or the model must be refined to allow for more reduction in uncertainty.

There are two primary reasons for the inherent uncertainty. First, the unresolved parameters combine to contribute significantly. Second, each of the resolved parameters still has uncertainty associated with it. As was mentioned in Section 3.7.5 on page 56, this study made a deliberate effort to avoid resolving parameters to an unrealistic degree.

Any form of the model, no matter how detailed, will contain uncertainty in a variety of forms: data, cutoff, aggregation, temporal, and geographic uncertainties. Details of these uncertainties can be found in Section 2.1.5 on page 23.

4.2.4 Weaknesses of Methodology

Various weakness were found with the specific LCA streamlining methodology pursued; the most significant of these are discussed below. Some of these weaknesses provide an opportunity for future work in this area and are discussed in that light in Section 6.3 on page 92.

The metric for contribution to uncertainty failed to identify a significant parameter

It was found that the partial rank correlation coefficients were ineffective when evaluating the significance of total engine mass. As previously mentioned, an exception to the PRCC selection criterion was made for engine mass; this parameter was resolved first, and proved to be the most significant parameter. This weakness is significant: PRCCs were unable to identify the most important parameter. Therefore, care should be taken when evaluating metrics for contribution to uncertainty.

No consideration for work involved in data collection

All resolution steps were considered equal in this work; therefore, it is very feasible that data that is difficult to collect would be prioritized over much more accessible data, even though the latter may in the end enable more resolution for the effort invested. Furthermore, these more time-intensive data collection steps could possibly be divided into smaller and more manageable pieces, each of which could have a measurable impact on resolution.

Little consideration for relative costs of parts

Very little cost data was available for this study: only the cost rankings of the parts were provided. One of the primary drawbacks of this limitation was that it was not possible to identify parts that had a lower energy use than their costs would predict. Identifying these outliers would contribute to a more accurate model.

Large portion of engine was approximated

Because of the large number of individual parts (> 1000) only a small fraction could be analyzed in detail. The largest parts were targeted, and therefore over half of the mass was modeled in detail. Because of helpful data on the composition of the entire engine, the remainder of the engine that was not modeled in detail had corresponding material composition information. Despite these measures, much of the details of the engine were modeled with high uncertainty. In Section 6.3.5 on page 94 one possible remedy is suggested: aggregating many small parts into larger “superparts” for more detailed analysis.

4.3 Recommended Streamlined Approach for Manufacturer

In this section we detail our specific recommendations to the manufacturer for performing streamlined life cycle assessments of their engines. These recommendations come from the lessons learned while producing this case study.

To improve the resolution of the model, it will be necessary to bring more engine components into the model beyond the 38 considered in detail in this study. Individually adding specific parts with compositions and processes that are energy-intensive may help; guidance may be needed to make this selection, preferably from a source very familiar with a specific engine model’s entire BOM. Another approach that may be very effective is the creation of aggregate parts based on part type. This concept is explored further in section 6.3.5 on page 94.

In a similar vein, internal expertise in life cycle assessment should be developed in the organization. On multiple occasions in this study detailed knowledge on both engine production and LCA was necessary. It should also be emphasized that the data gathering process should be interdisciplinary and involve multiple functional groups inside the organization. As an example, some of the most influential data for this study - the aggregate material composition of the target engine - came from the purchasing department, far outside the design engineering group where most of the data was gathered.

Based on the results in this study, it may be possible to forgo detailed modeling of supply chain transportation for individual parts, as the corresponding reduction in uncertainty is relatively small. It may, however, be very beneficial to collect aggregate and general data on supply chain transportation. For example, the travel could be modeled loosely by the countries involved rather than the specific locations. The model could then incorporate tighter bounds than crude and general approximations for all parts.

Details of engine material composition can also be limited to the dominant materials. Over 90% of the remainder of the engine was aluminum and low alloy steel. These mass parameters indeed had high leverage; however, the rest of the compositional parameters lacked significant leverage.

Chapter 5

Additional Engine Production Assessments

While the development of a streamlined LCA methodology for diesel engines was the primary focus of this research, other assessments of diesel engine production were also done carried out. Each of these additional assessments was done in collaboration with the same engine manufacture. The analyses range from the specific comparison of two camshaft designs to a broad assessment of energy use at the factory scale.

5.1 Streamlined LCA Results for Case Study Engine Production

As a natural extension of the streamlining experiments, a complete streamlined LCA for the case study engine was produced with five impact metrics: embodied energy (discussed in Section 3.2.2 on page 36) and four additional metrics, described below. This LCA utilized the case study model and LCI data from the Ecoinvent database.

5.1.1 Impact Metrics and Scope of LCA

- Greenhouse gas emissions. Greenhouse gas emissions are commonly expressed as the global warming potential of the substances involved. Global warming potential is a measure of the radiative forcing of a gas in the atmosphere. Global warming potential and greenhouse gas emissions are expressed in CO_2

equivalence, which normalizes a substance's global warming potential by that of carbon dioxide. [58]

- Acidification potential. Acidification potential is a measure of a substance's disposition or potential to release H^+ ions. The measure of acidification used here is SO_2 equivalence: a substance's acidification potential per mass normalized by that of sulfur dioxide. [59, 60, 61]

- Eutrophication potential. Eutrophication potential is a measure of a substance's potential to cause over-fertilization of soil and/or water, and thereby increasing the growth of biomass and potentially decreasing biodiversity. The measure of eutrophication potential used here is PO_4 equivalence, which normalizes by the eutrophication potential of PO_4^{3-}. [59, 60, 62]

- Embodied water. Embodied water, like embodied energy, measures the water used to produce a substance. The Ecoinvent database considered water depletion from rivers, lakes, wells, and all other sources of fresh water. Embodied water is expressed as the total volume or mass of the fresh water used. [63]

5.1.2 Embodied Energy vs. Other Environmental Impact Metrics

The success of embodied energy as an environmental proxy metric can be tested using the results of this streamlined LCA. If the relative impact of the parts is similar under two different metrics, the metrics may be appropriate substitutes for each other in this study. Table 5.1 on page 73 compares each metric to embodied energy: greenhouse gas emissions, acidification potential, eutrophication potential, and embodied water use.

Embodied energy appears to be a very good proxy for greenhouse gas emissions. There is only one noticeable outlier in this study, the valve cover. The valve cover is unique among the parts studied because it is comprised completely of plastic (polyethylene terephthalate, PET). The relative impact of this plastic part is higher in embodied energy than greenhouse gas emissions.

Acidification potential and eutrophication potential are less easily represented by embodied energy. The plastic valve cover is again an outlier when these two metrics are compared with energy, and again the relative impact is less for these metrics than for energy. Other outliers emerge as well. The most salient outlier is the nitrogen sensor, whose relative impact is much greater in these two metrics than in embodied energy. This sensor is comprised of copper and tin, both of which contributed to the uncorrelated metrics.

Water use shows perhaps the weakest correlation with energy use. A general correlation is visible, but there are many outliers. The valve cover and the nitrogen sensor are again two of the most significant outliers. Also notable is the exhaust transfer tube. The transfer tube's metal body is unremarkable in relation to the other parts, but it is coated in an acrylic resin.

For more detailed results of this LCA, please see Table B.8 on page 122 and Table B.9 on page 123, both in the Appendix.

5.1.3 Summary of LCA Results

Figures 5.2 and 5.3 summarize the results of the LCA in terms of embodied energy. Figure 5.2 groups embodied energy by category. The most significant result is the importance of material embodied energy, which accounts for a full two-thirds of embodied energy.

Figure 5.3 groups embodied energy by engine part. The largest parts on the engine predictably have the largest embodied energies: the cylinder block, the crankshaft, and the cylinder head. However, the nine most energy-intensive parts shown in the figure only account for a little more than half of the total embodied energy.

5.2 Camshaft Impact Comparison

The engine's camshaft was studied in more detail to provide a detailed LCA case study. The primary impetus for this research task was comparing two competing camshaft production processes. There are a variety of production techniques available for manufacturing camshafts, detailed below.

5.2.1 Camshaft Manufacturing Methods

Cast Camshafts Camshafts are commonly cast from either steel or iron alloys. The resulting casting then undergoes turning and grinding on all contact surfaces,

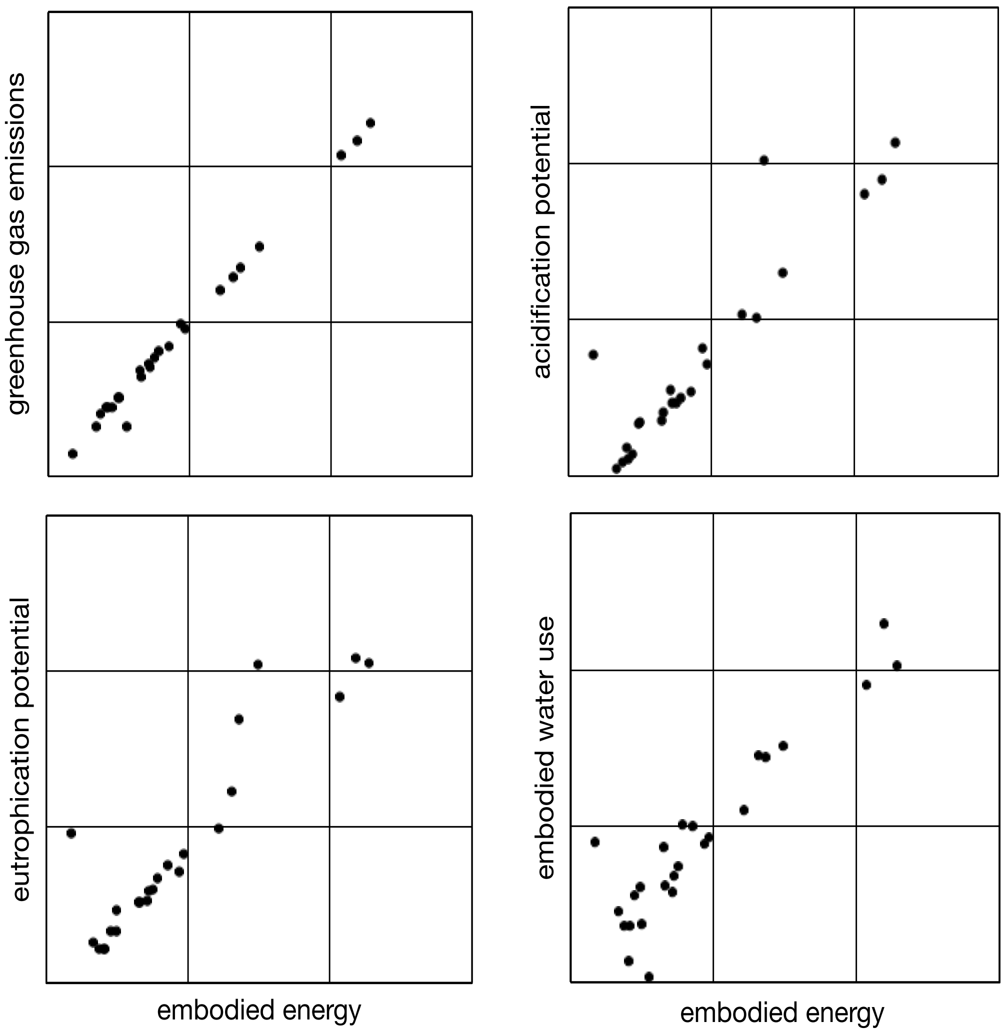


Figure 5.1 – Embodied energy compared with other environmental metrics for 25 engine parts. The relative impact of the 25 most energy-intensive parts is plotted on both axes of the log-log plot. The x-axes are the relative impact measured in embodied energy; the y-axes are the other environmental impacts studied. If two impacts are highly correlated they may be adequate proxies for each other. For example, this is the case for embodied energy and greenhouse gas emissions. The details of these plots are given in Table B.9 on page 123.

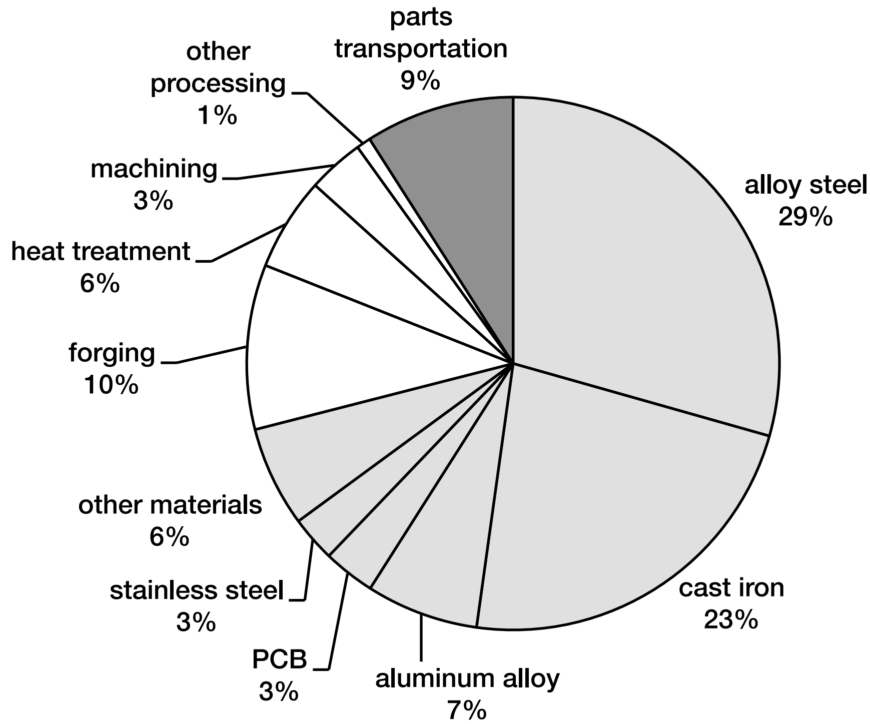


Figure 5.2 – Energy use by category in case study LCA. The categories are grouped into materials (light gray), processes (white), and transportation of parts (dark gray).

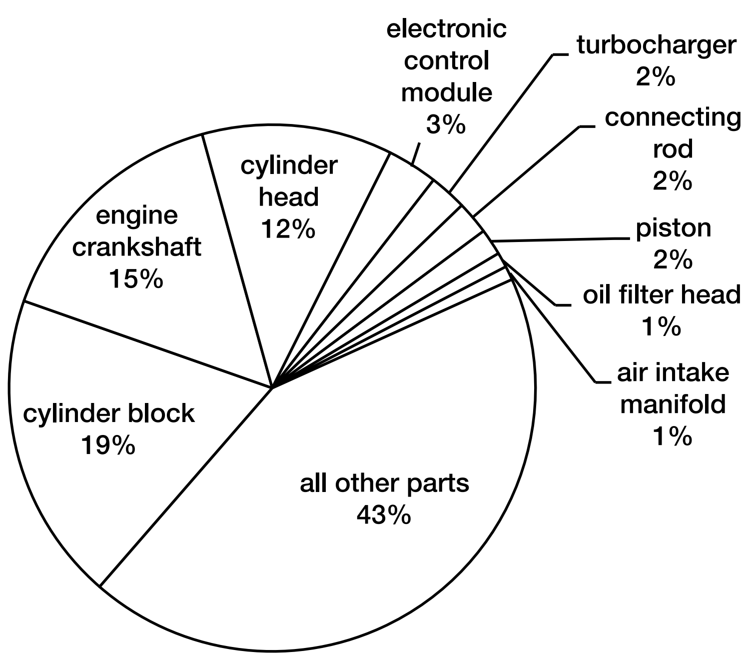


Figure 5.3 – Embodied energy by engine part in case study LCA.

such as the lobe and bearing surfaces. In general, cast parts are economical to produce and relatively strong.

Forged Camshafts Forged camshafts are hot-forged from a single casting of steel or iron. The forging surface can then be machined on contact surfaces in a similar manner as a cast camshaft. Forgings may also be produced with high quality surface finishes, therefore obviating further machining. Forged parts are commonly stronger per unit mass than comparable cast parts.

Machined Camshafts Machined camshafts are produced from bar stock (usually steel). The shape of the camshaft is produced through milling, turning, and grinding. The camshaft is commonly heat treated afterwards for increased strength.

Assembled Camshafts Assembled camshafts are a relativity recent innovation. This design is a conglomeration of a variety of parts and materials. The center shaft may be solid or hollow. The cam lobes may be either sintered or forged and usually have machined surfaces. The lobes are assembled onto the shaft, typically through thermal expansion. The assembled camshaft's primary advantages are the novel materials available and potential cost savings for low volume production runs.

5.2.2 Direct Data Measurements

Experimental data was gathered for a detailed analysis of the camshaft production. Measurements were taken in 2012 at a manufacturing facility in New York. These camshafts were produced by machining. The camshaft production line consisted of 45 separate production machines performing 15 discrete processing steps. The sole energy inputs to all but one process was grid electricity. A single process, the draw oven heat treatment, consumed natural gas.

All electrical equipment measured was powered by a 480-volt 3-phase supply. Electricity consumption was measured both at substations and individual machines, depending on the power rating and the wiring configuration. Most processes had cycle times of less than ten minutes; multiple cycle times were evaluated for each machine, with a target of 50 cycles. Continuous power consumption measurements were logged using an AEMC 3945-B three-phase power quality analyzer. Average measurements for each of the three phases were logged every second.

Processing Step	Energy per part (MJ)	% of total
Lobe Mill	124.3	31%
Heat Treatment	82.8	20%
Bearing Grinder	75.7	19%
Lobe Grinder	57.8	14%
Lathe	42.2	10%

Table 5.1 – Energy use for major processes for machined camshaft production, measured in MJ per part. All electricity measurements were converted to primary fuel energy. These five processes used 94% of the total production energy for machined camshafts.

The energy used for each processing step was calculated using the instantaneous power consumption integrated over the duration of the step. Idle power for all machines was also measured. Most of the process machines remained powered while not in use. The average energy use per camshaft produced therefore varied noticeably with the daily production volume.

For this study, primary fuel energy was the measure of energy used. Primary fuel energy traces electricity use back to the fuels burned and allows for the comparison of both electricity and natural gas consumption. The conversion from electricity energy E_e to primary fuel energy E_p was performed using a conversion efficiency factor for the United States, 38.7%, from the International Energy Agency [64]:

$$E_p = \frac{E_e}{0.387} = 2.58E_e$$

5.2.3 Measurement Results

It was found that out of the 16 processes, a few used the vast majority of the energy. Two processes, lobe milling and heat treatment, used over half of the total primary energy. The top five processes used 94% of the total. See Table 5.1 for details on the energy used in these steps. All processes and their energy usage are depicted in a Pareto chart in Figure 5.4 on the facing page.

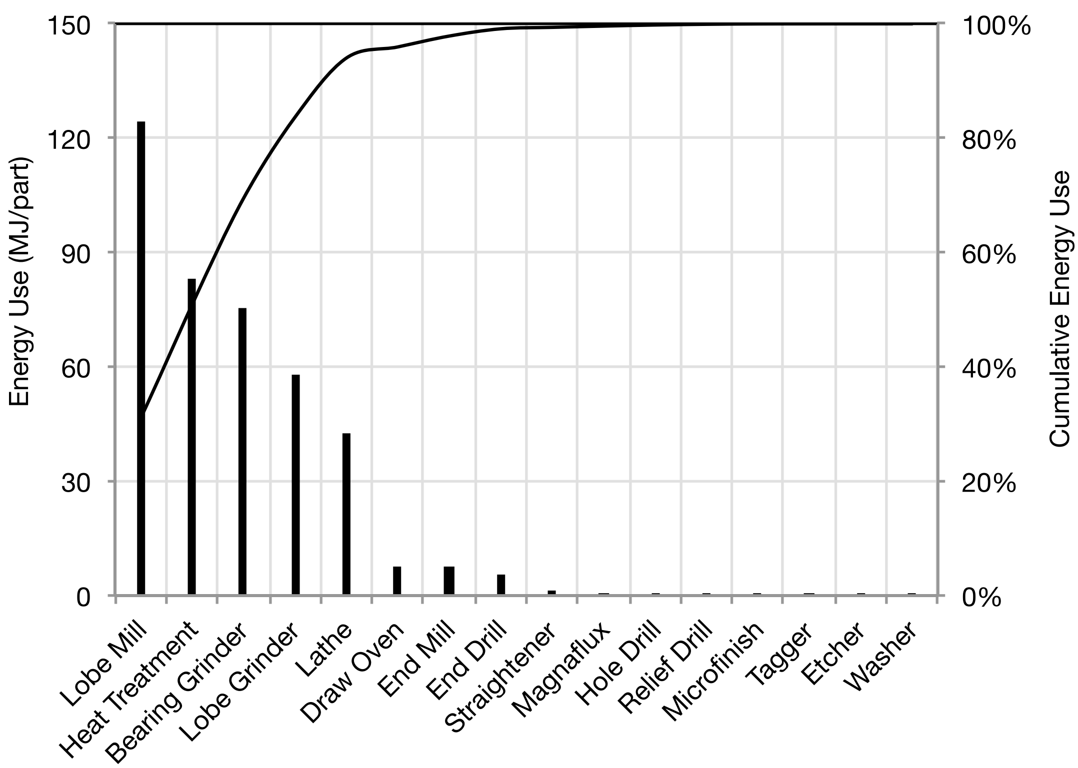


Figure 5.4 – Machined camshaft: Pareto chart of energy use over the 16 processing steps. Note the large share that the first five processes contribute to the total. All energy is measured in primary fuel energy.

5.3 Manufacturing Plant Energy Usage

Energy usage data was also collected for entire engine production facilities. The primary impetus was finding the strength of the relationship between plant-wide energy consumption and production. Other similar studies have shown a weak connection between production variation and overall energy consumption [65, 66].

5.3.1 Facilities Profiles and Data Collection

The energy use of two different factories was investigated. The first plant assembled large diesel truck engines (FHWA Class 6 - 8 [29]); the primary function of the plant was engine assembly, accompanied by some other processing, such as machining and heat treatment. The second facility produced diesel turbochargers. The operations in this facility primarily consisted of turbocharger assembly.

Both electricity and natural gas use were considered as energy inputs into the plants. The energy use data was gathered from monthly plant records over the course of two years: January 2010 to December 2011. This data was then compared to production records for the same periods.

In this section "unit" or "product" will refer to the specific outputs of the different plants - engines and turbochargers, respectively. The primary metric studied was aggregate energy use per unit produced. The significance of this energy use is always relative to the base load. This base load represents the energy used to power all facility-wide devices: air conditioning and heating units, plant lighting, and office equipment. It also represents any idle energy consumption by the production machines themselves. Above the baseline, a trend may emerge, representing the incremental energy consumption of the entire facility as production volume varies over time.

5.3.2 Engine Assembly Facility

The engine assembly facility had a very pronounced base load electricity use, which accounted for 56% to 83% of the total electricity used in a month. The incremental production electricity was clearly visible as well, increasing roughly linearly. Figure 5.5 on the next page illustrates total energy use versus production volume; Figure 5.6 on page 80 demonstrates the variation in electricity use per engine. As production increases, the plant's apparent energy efficiency rises, as each engine requires less

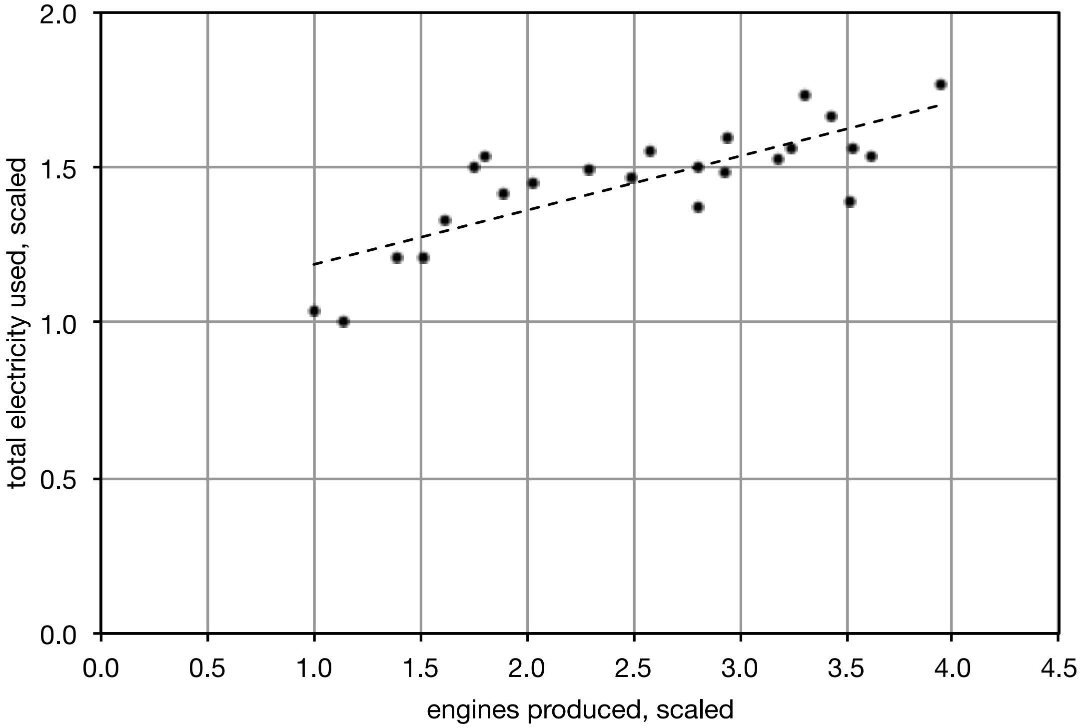


Figure 5.5 – Engine plant electricity usage vs. units produced. The electricity and units produced have been scaled linearly to protect propriety production volume information.

electricity. The slope of the curve in Figure 5.5 is the minimum electricity use per engine, 193 kw · h; this is the asymptotic limit of energy used per unit as production volume grows.

Natural gas use shows no correlation with engine production whatsoever, as Figure 5.7 on page 81 shows. This leads to the conclusion that the natural gas heat treatment processes were overshadowed by natural gas use for building heating. This hypothesis was further tested by comparing natural gas consumption to historical weather records [67]. Figure 5.8 on page 82 shows a clear inverse relationship between average monthly temperature and natural gas use.

5.3.3 Turbocharger Facility

Electricity use at the turbocharger facility showed a much higher dependence on production volume. In some cases incremental load was five times larger than the base load. Figures 5.9 on page 83 and 5.10 on page 84 illustrate total electricity use

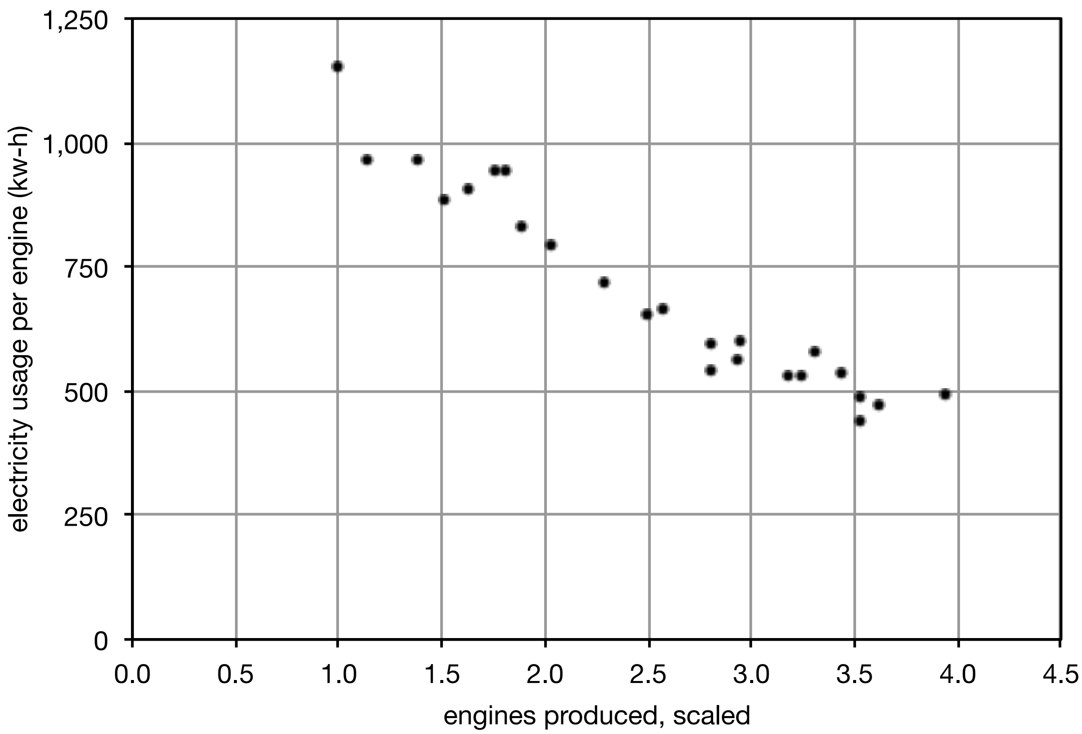


Figure 5.6 – Electricity used per engine vs. units produced at the engine plant. Production volume has been scaled in the same manner as in Figure 5.5;

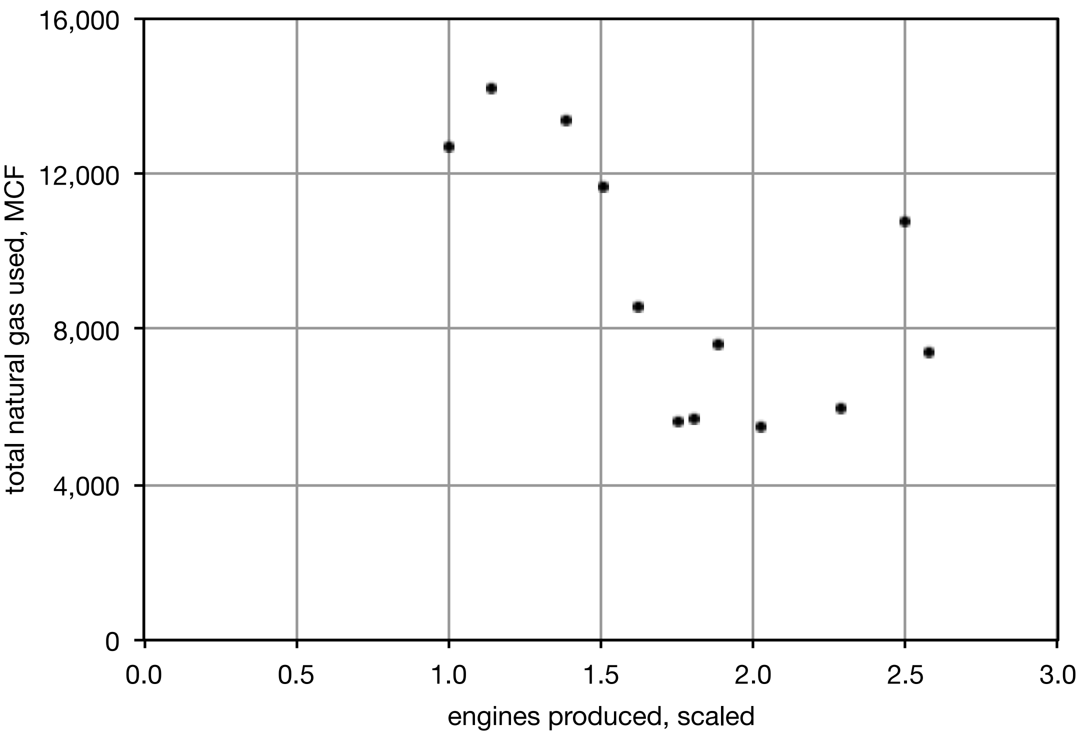


Figure 5.7 – Engine plant natural gas usage vs. units produced. Production volume has been scaled in the same manner as in Figure 5.5.

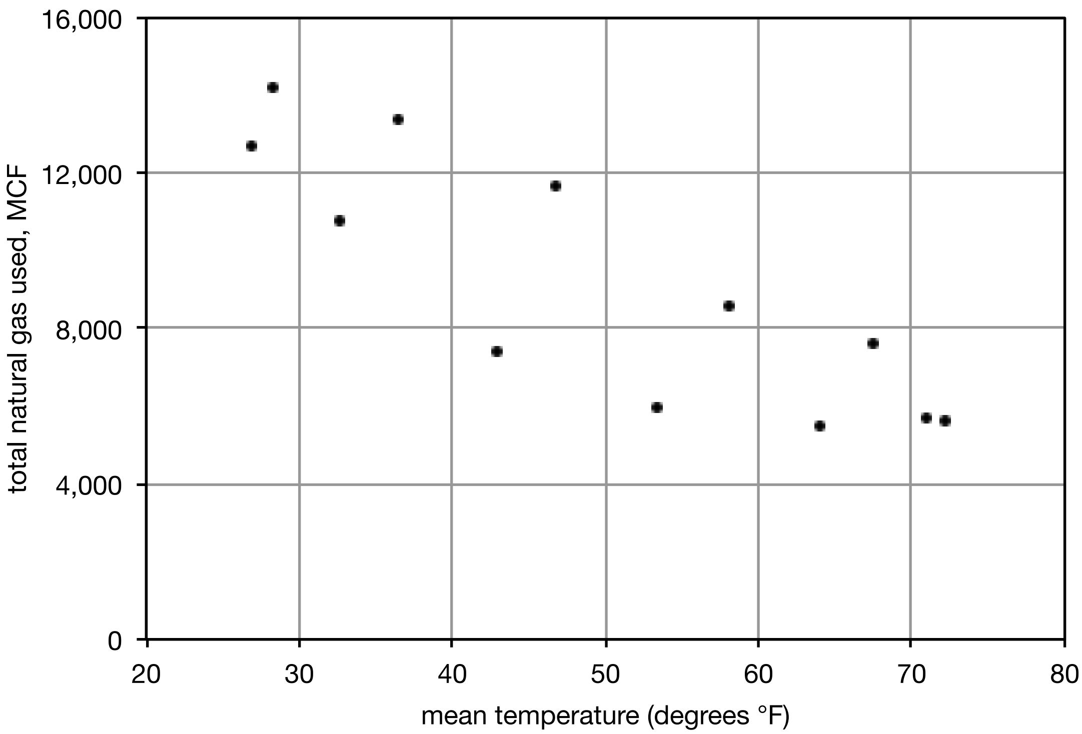


Figure 5.8 – Engine plant natural gas usage vs. mean temperature. Production volume has been scaled in the same manner as in Figure 5.5.

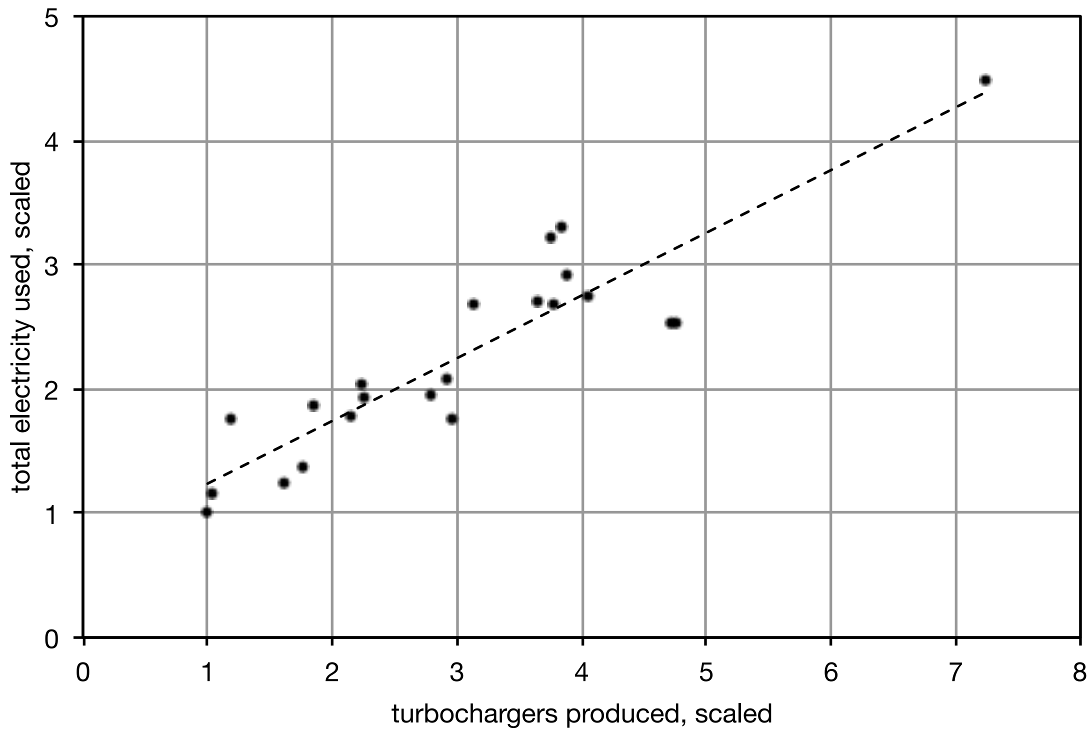


Figure 5.9 – Turbocharger plant electricity usage vs. units produced. The electricity and units produced have been scaled linearly to protect propriety production volume information.

and electricity use per unit produced. Each unit produced uses approximately 11 kw-h of electricity.

5.4 Diesel Engine Composition Across Application Areas

The final analysis of energy consumption in diesel engine manufacturing focused specifically on embodied energy of materials. Our collaborating manufacturer produces engines of varied size and power output. These engines correspond to various application areas, such as commercial and industrial power generators, truck engines, and marine engines. The material composition of these types of engines varies across the application areas, and therefore the energy invested in each engine varies as well.

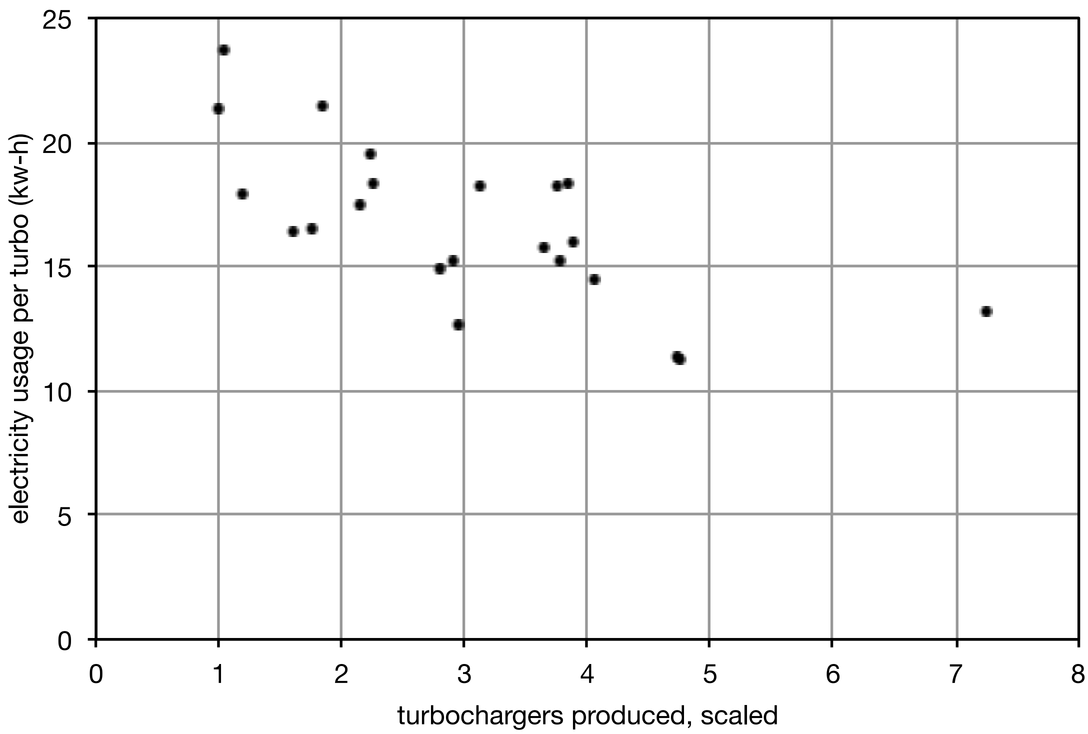


Figure 5.10 – Electricity used per turbocharger vs. units produced. Production volume has been scaled in the same manner as in Figure 5.9.

5.4.1 Importance of Material Embodied Energy

In the LCA of the case study engine in Section 5.1.3, the importance of materials in overall energy use is clear. Materials account for approximately two-thirds of the total energy used to produce that engine. Because of this importance, restricting a study to the material composition of various engines should still lead to relevant results.

5.4.2 Engine and Material Background Information

The data for this section was acquired from purchasing records in the same manner as the material composition for the case study engine (discussed in Section 3.6.4 on page 49). Twenty-seven different engines were studied. The engines ranged in mass from 190 kg to 13,500 kg and in engine displacement from 2.8 L to 91 L. As evidenced by the range of mass and displacement, the engines spanned many application areas. The engines can be grouped into four general categories, listed below. The categories delineate between on- and off-highway engines. On-highway engines are used in vehicles that travel on public roads and must therefore meet specific emissions standards. In the United States, these standards are set by the Environmental Protection Agency [68]. Off-highway vehicles and engines are not subjected to the same standards.

- On-highway, medium-duty - The engines range in displacement from 2.8 L to 6.7 L. Example applications include delivery trucks and small buses.
- On-highway, heavy-duty - The engines range in displacement from 8 L to 15 L. Examples of uses include fire trucks, buses, and semi-trailer trucks.
- Off-highway - The engines range in displacement from 15 L to 28 L. Example applications include bulldozers and other mining vehicles.
- Power generation - The engines range in displacement from 19 L to 91 L. Examples of uses include locomotives, ships, and industrial power generation.

The composition of the engines was broken into 18 separate materials; details of these materials, as well as data sources used, can be found in Table B.10 on page 126. While these materials are not all inclusive, in many cases they account for 99% or more of the total engine mass.

5.4.3 Material Composition of Various Engine Types

Below are general observations about the material composition of the engines. The material composition varied considerably across the engine lines, but some materials were common to all.

- High scrap steel content. All engines - regardless of type, size, or other materials used - were composed of 41% to 47% scrap steel by mass.

- Few extremely energy-intensive materials. The four most energy-intensive materials in this audit were nickel, molybdenum, tin, and titanium[1]. The highest nickel content among the engines was 1.0% by mass; most were less than 0.8%. The other three energy-intensive materials were individually at most 0.36% of an engine's mass.

- Narrow range of pig iron and low-alloy steel usage. Like scrap steel, both pig iron and low-alloy steel were abundant in all the engines with narrow ranges. By mass, all engines contained between 18% and 20% pig iron and between 26% and 27% low-alloy steel.

5.4.4 Embodied Energy Findings

The composition of each engine was used to determine the approximate embodied energy in the materials. The most notable result is the strong linear relationship between engine mass and embodied energy across all engines analyzed (See Figure 5.11 on the facing page). Stated another way, the material embodied energy per unit mass ($^{\mathrm{MJ}}/_{\mathrm{kg}}$) for all engines was similar; Figure 5.12 on the next page illustrates the finding that all engines have an average embodied energy of 17 to 20 $^{\mathrm{MJ}}/_{\mathrm{kg}}$. The smaller engines have larger embodied energies, from 18.8 to 19.8 $^{\mathrm{MJ}}/_{\mathrm{kg}}$; the largest engines have embodied energy values of about 17.2 $^{\mathrm{MJ}}/_{\mathrm{kg}}$.

[1]Approximate embodied energy of these materials:
nickel, 142 $^{\mathrm{MJ}}/_{\mathrm{kg}}$
molybdenum, 151 $^{\mathrm{MJ}}/_{\mathrm{kg}}$
tin 321, $^{\mathrm{MJ}}/_{\mathrm{kg}}$
titanium, 670 $^{\mathrm{MJ}}/_{\mathrm{kg}}$
Please see Table B.10 on page 126 in the Appendix for more details on these materials.

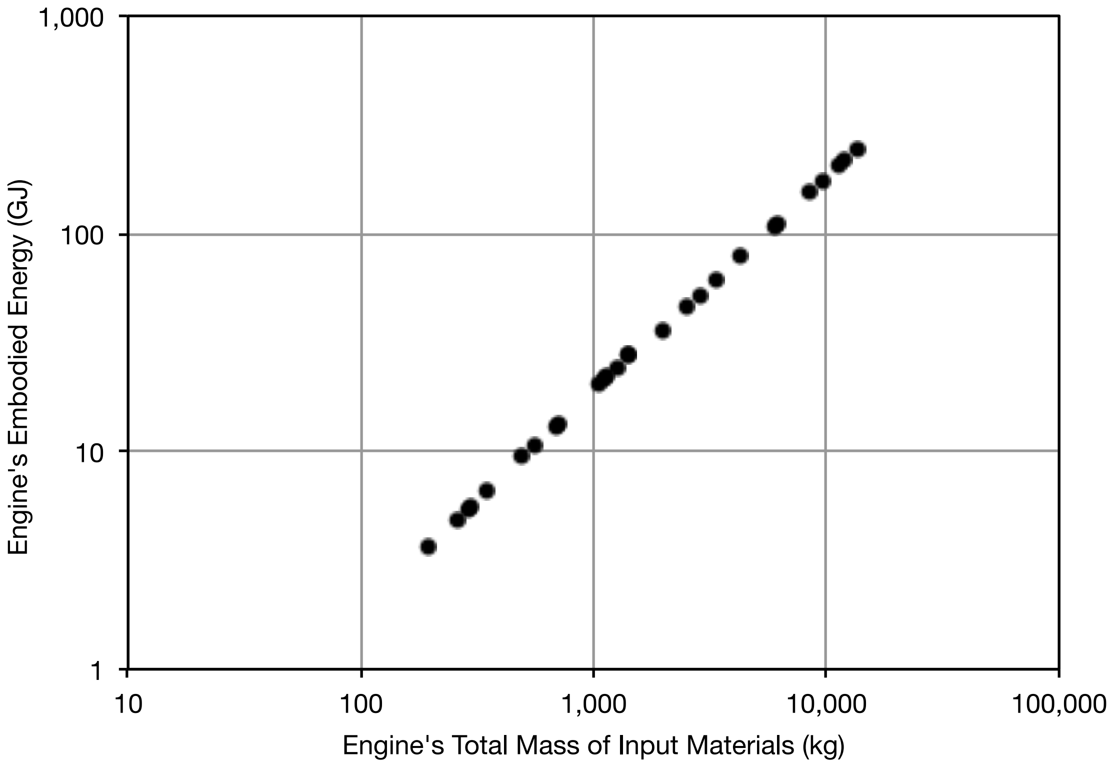


Figure 5.11 – Embodied energies of engines vs. total engine mass, log-log scales. Note the strong linear relationship.

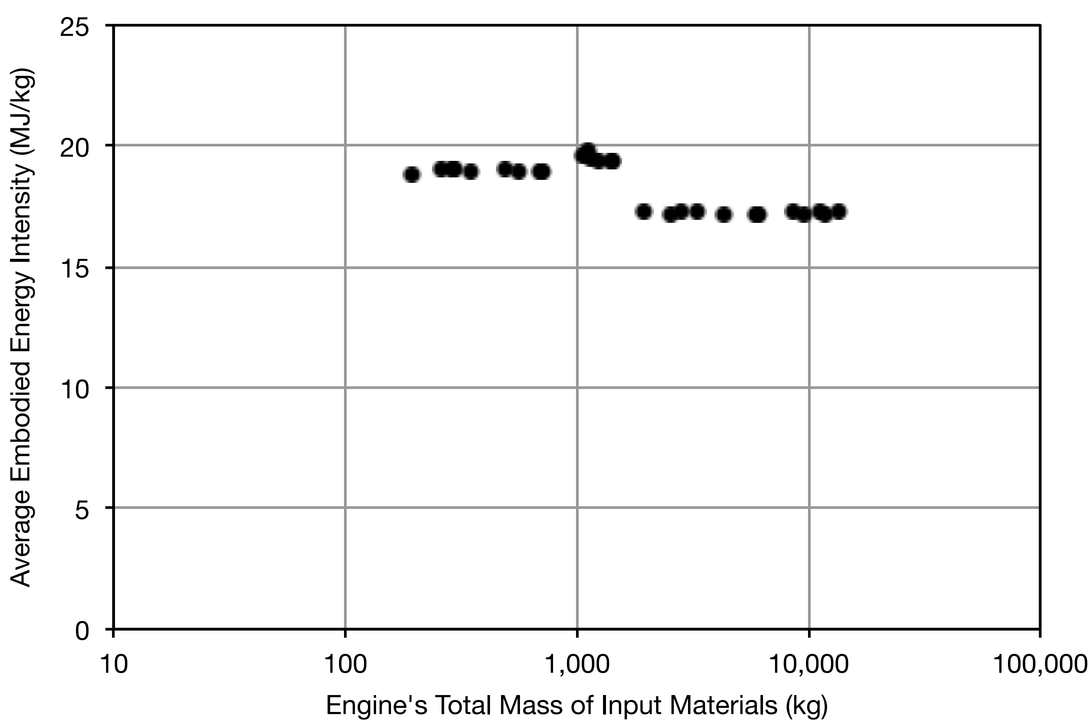


Figure 5.12 – Average embodied energy intensities of engines vs. total engine mass. There is very little change in the energy intensity over a 100-fold change in engine mass.

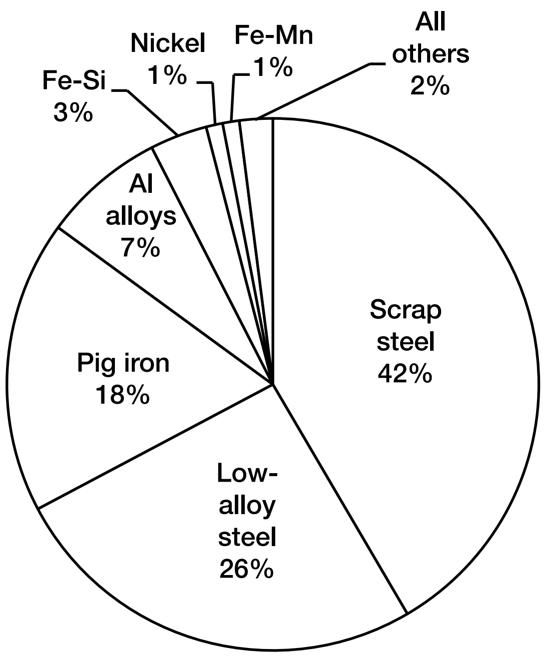
To explore the variations in composition and energy intensity between engines of different sizes, two representative engines were chosen: an 11 L on-highway engine and a 19 L off-highway engine. For brevity and clarity, these engines will be referred to as "small" and "large," respectively. These two engines represent the largest (the small engine) and smallest (the large engine) embodied energy intensity. Table 5.2 summarizes the two engines.

Engine	Displacement	Category	Mass	Embodied Energy Intensity
"Small"	11 L	on-highway	1,104 kg	19.8 MJ/kg
"Large"	19 L	off-highway	1,973 kg	17.3 MJ/kg

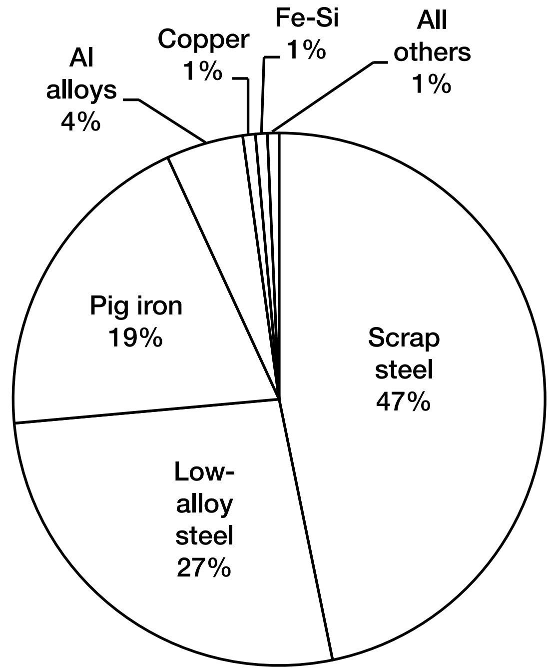
Table 5.2 – Details of representative engines.

Figures 5.13 a-d on the next page explore the differences in composition between the small and large engine. The compositions by mass of the small and large engines are shown in 5.13a and 5.13b, respectively. The large engine contains a larger portion of low energy intensity materials: scrap steel, low-alloy steel, and pig iron; these three materials account for 93% of the large engine's mass and 86% of the small engine's mass. This is a major factor in the lower energy intensity of the large engine.

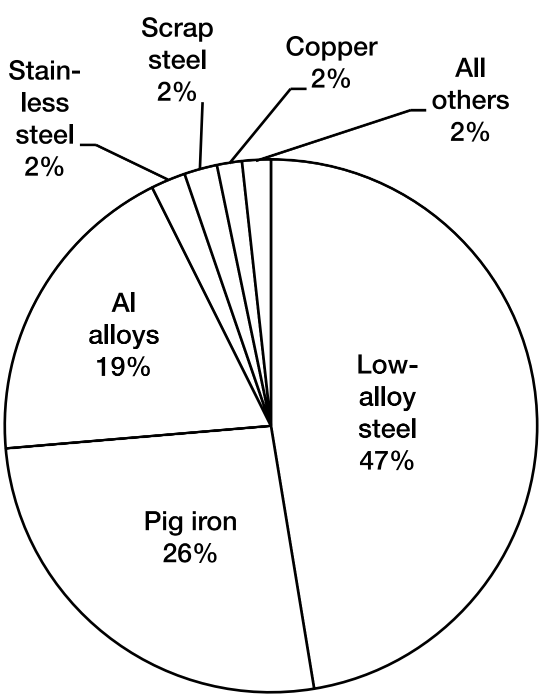
The material embodied energy breakdowns for the small and large engines are shown in 5.13c and 5.13d, respectively. The most striking feature of these figures is the extremely small embodied energy contribution of scrap steel; although it accounts for almost half of the mass of the engines, it contributes only about 2% of the embodied energy.



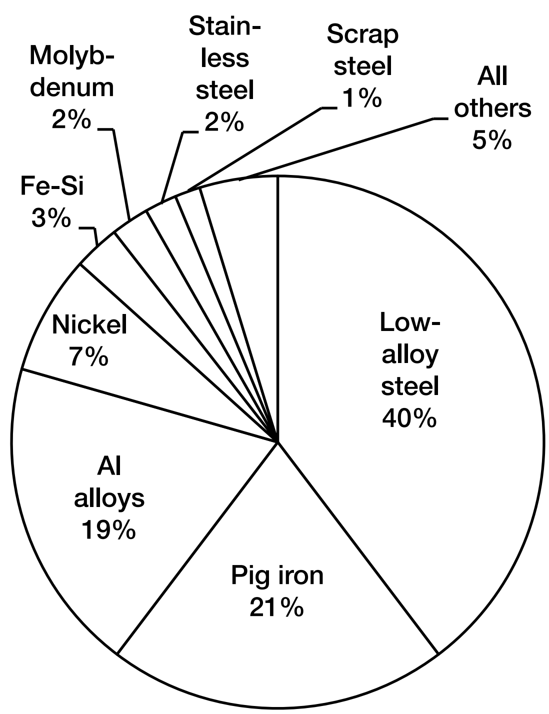
(a) mass, small engine



(b) mass, large engine



(c) embodied energy, small engine



(d) embodied energy, large engine

Figure 5.13 – Material mass (top row) and embodied energy (bottom row) of a representative small (left column) and large engine (right column). See Section 5.4.4 for details.

Chapter 6

Conclusion

6.1 Research Background

This research addresses the complexity and burden of LCA by highlighting important product parameters and thereby guiding data collection. This work may be categorized as LCA streamlining, which is the process of reducing the necessary effort to produce acceptable analyses. While many methods of streamlining exist, most select *a priori* which activities and parameters to evaluate. This selection relies very heavily on the researchers' intuition, and may ignore large impacts. The reduction in effort expended is therefore often accompanied by a commensurate increase in the uncertainty of the results.

6.2 Contribution of Work

This research project performed a streamlined LCA case study with a diesel engine manufacturer. A specific diesel engine was selected: a large 15-liter on-highway truck engine. Embodied energy was selected as the primary environmental impact metric. A complex model of its energy use in the production and manufacturing life cycles was created. The model consisted of 184 total parameters describing various aspects of production energy use: part mass, part composition, material energy intensity, process energy intensity, transportation distance, and transportation mode. Each parameter was assigned a generous range, derived from readily available data. The model was then simulated repeatedly and the parameters were randomly assigned, producing a range of possible model outputs.

Each parameter's contribution to model uncertainty (or variance) was approximated using partial rank correlation coefficients (PRCCs). PRCCs are designed for nonlinear models, but assume a monotonic relationship between the model inputs (engine parameters) and outputs (energy use). This assumption was shown to be effective for most parameters. The variance contribution of one parameter - total engine mass - was found to be noticeably underestimated by PRCCs. This finding led to a simple heuristic to prioritize this parameter above all others. The model simulations were then rerun.

After each simulation, the parameter with the most significant contribution to uncertainty (as judged by PRCC) was resolved. The resulting decrease in model uncertainty was then evaluated. Total model uncertainty was quantified by a custom test and metric: the self-test and the False Error Rate (FSR), respectively. The self-test compares two hypothetical engines with embodied energy values differing by 10%. This is accomplished by creating an identical impact distribution (B) shifted from the original (A) by +10%, such that $E(X_B) > E(X_A)$. The FSR is then the rate of confusion between the two engines, the instances that falsely conclude that $X_B < X_A$. The maximum FSR possible is 50%. Initially the model FSR was 21%. After 20 parameters were resolved it had decreased to 6.1%; it finally reached a steady state at 5.8% after 39 parameters were resolved. The subsequent resolution of any parameters - indeed, all of the remaining parameters - did not affect the uncertainty.

Many conclusions can be drawn from this case study. Most saliently, few parameters in the engine energy use model were significant; therefore, these parameters should be targeted for data collection and resolution. These parameters included overall approximate engine mass, approximate composition, and the masses of some parts. For certain large and expensive parts energy intensity of materials and processing were also significant. A few smaller and less expensive parts affected the model, but typically only via their material energy intensity.

The transportation parameters for individual parts were not selected for resolution, and did not contribute to model uncertainty in a discernible manner. This is a significant insight: total transportation accounts for a significant portion - roughly 9% - of the total energy use in this study's boundaries, yet its constituent parameters have no appreciable leverage in the model.

Finally, the model has significant inherent and unavoidable uncertainty. This has significant consequences for data collection. The inherent uncertainty in the formulation dominated most of the parameters in the model, and therefore less than 45 were significant enough to affect overall uncertainty. Data should therefore not be collected on the insignificant parameters.

6.3 Suggestions for Future Work

Following this research, there are multiple possibilities for continued work. Below are suggested projects that may be particularly relevant to this work.

6.3.1 Data Collection Cost Optimization

In this research and in similar work, the resolution of each parameter is treated with equal weight. There is no consideration of the amount of investment needed to resolve different parameters. Work could be done to evaluate the relative difficulty and cost associated with resolving different product attributes (e.g. part mass, part composition, material energy intensity, distance transported, and mode of transportation). An optimization problem could then be formulated to achieve the desired resolution of the LCA for the lowest investment cost.

It may also be possible to integrate into the optimization problem the interaction between data-gathering steps. For example, gathering process data for the engine block may significantly lower the burden in gathering data for the fuel pump housing, if both are sourced from the same supplier. This interaction between steps may make it advantageous to gather multiple disparate parameter data in one task. The optimization problem created by interdependence would be similar to linear optimization problems in open pit mining.[1]

6.3.2 Streamlining for Multiple Environmental Impact Metrics

While one environmental impact - energy use - served as the sole metric for this study, the resulting streamlined LCA could be completed using any metrics. However, this approach has one clear downside: the uncertainty associated with other metrics, such as water use, may be very different than that of energy use. The resulting uncertainty

[1] In open pit mining optimization, removing obstructions to certain caches of ore may make it advantageous to also mine other caches (that would not have been otherwise feasible to mine). For examples of optimization problems in open pit mining, See Espinoza et al. [69]

of other metrics may then be unacceptably high. Future research could be dedicated to streamlining over multiple environmental impact metrics. Compromises would have to be made so that all metrics in question were reduced to an acceptable level of uncertainty.

6.3.3 Staged Resolution of Parameters

Another opportunity for improvement can be found in the types of uncertainty targeted in the resolution step. Resolution is not a simple or uniform process. Various types of uncertainty could be resolved (e.g. temporal uncertainty, geographic uncertainty, and process uncertainty). Therefore the resolution of a single parameter could actually be divided into multiple stages.

As an illustration, consider the model parameter for the energy intensity of crankshaft forging. Assume that based on contribution to model variance, it is determined that this parameter should be resolved. Rather than resolve all available types of uncertainty in one laborious data-gathering task, the resolution could be staged. First, the process uncertainty could be reduced by researching the specific forging process utilized. After this resolution, the model could again be simulated. If the forging of the engine crankshaft no longer contributed most significantly to model variance, it would not have to be resolved further in this step. If resolution was still desired, the temporal uncertainty could be resolved; the most recent data for the specific forging process would be needed. Other possible stages are the specific geographic area of the forging and the specific forging plant and process.

This staging strategy could also be used to address multiple suppliers. Rather than simultaneously gathering data from all suppliers of a specific part, a single supplier - possibly the most commonly used supplier for this part - could be targeted for resolution.

Staged resolution could also greatly benefit from the aforementioned cost optimization of data collection, as some stages may be much more difficult to resolve than others. Patanavanich [22] explores a type of staged resolution applied specifically to materials.[2]

[2]See Section 2.3.3 on page 29 for a brief description of this work.

6.3.4 Accounting for Cost Anomalies in the BOM

The bill of materials detailed certain parts that are much more expensive than their mass or material would predict. Examples of anomalous parts are the fuel injectors: the six injectors, with an approximate total mass of less than 1 kg, were in the top five most expensive parts (if all six are treated *en masse*) along with much more massive parts: the cylinder block (>300 kg), the cylinder head (>200 kg), and the engine crankshaft (>100 kg). The fuel injector material is tool steel; while relatively high, the material cost of tool steel cannot explain the large cost of the injectors. Energy intensive precision machining most likely accounts for a sizable portion of the injector cost. Anomalies such as these could be pursued to produce a more accurate model of energy use.

To identify these outliers, a hybrid LCA approach could be used. Only crude measures of cost of goods sold (COGS) was made available for this study in the form of the relative cost rankings of the parts. If exact part cost data was made available, one method of identifying cost outliers would be to use a course economic input-output LCA model, such as the EIO-LCA model [16]. The engine model's estimation for energy used to produce a part could be compared with the EIO-LCA estimation; those parts that showed the greatest relative underestimation could be selected for further study.

6.3.5 Part Aggregation

The single improvement that would most help this research would be part aggregation. As discussed in this thesis, the investigated engine contains over 1,000 individual parts. Many of these are small parts with masses on the order of 0.1 kg; however, their aggregate impact could be significant. As an illustration, there are 172 fasteners in the bill of materials. Even if these fasteners had an average mass of only 0.05 kg (50 g), their total mass of 8.6 kg would be greater than many of the 38 most expensive parts, which were culled for detailed analysis in this study. These fasteners could be combined to form a part aggregation, or a "superpart," that could be analyzed along with the other parts of the engine, thereby decreasing the mass of the remaining parts and increasing the possible resolution of the model.

The implementation of aggregate parts could take various forms: it might be thorough or it may simply rely on intuition and generous uncertainty bounds, as in the example above. Even in the latter case, this aggregation may provide more insight into the impact of the product under study.

Bibliography

[1] H. Wenzel. Application dependency of LCA methodology: Key variables and their mode of influencing the method. *The International Journal of Life Cycle Assessment*, 3(5):281–288, 1998.

[2] E. Hochschorner and G. Finnveden. Evaluation of two simplified life cycle assessment methods. *The International Journal of Life Cycle Assessment*, 8(3):119–128, 2003.

[3] A. Jensen, L. Hoffman, B. Møller, and A. Schmidt. *Life Cycle Assessment (LCA): A Guide to Approaches, Experiences and Information Sources*. Environmental Issue Report. European Environment Agency, 6 edition, 1998.

[4] J. Dettling and M. Margni. Comparative environmental life cycle assessment of hand drying systems: The XLERATOR hand dryer, conventional hand dryers and paper towel systems. Technical report, Quantis US, 2009.

[5] Environmental Label Jury. The Blue Angel - eco-label with brand character, 2013.

[6] Environmental Protection Agency. Life cycle assessment: Inventory guidelines and principles. Technical report, Office of Research and Development, Cincinnati, Ohio, USA, 1993.

[7] B. Maurice, R. Frischknecht, V. Coelho-Schwirtz, and K. Hungerbihler. Uncertainty analysis in life cycle inventory. application to the production of electricity with French coal power plants. *Journal of Cleaner Production*, 8(2):95–108, 2000.

[8] J. McAuley. Global sustainability and key needs in future automotive design. *Environmental Science & Technology*, 37(23):5414–5416, 2003.

[9] M. Ashby. *Materials and the Environment*. Butterworth-Heinemann, Oxford, UK, 2009.

[10] J. Sullivan, R. Williams, S. Yester, and E. Cobas-Flores. Life cycle inventory of a generic U.S. family sedan overview of results USCAR AMP project. SAE Technical Paper 982160, SAE, 1998.

[11] T. Gutowski, S. Sahni, A. Boustani, and S. Graves. Remanufacturing and energy savings. *Environmental Science & Technology*, 45:4540–4547, 2011.

[12] International Organization for Standardization. Environmental management - life cycle assessment - principles and framework: ISO 14040: 2006. Technical report, International Organization for Standardization, 2006.

[13] International Organization for Standardization. Environmental management - life cycle assessment - requirements and guidelines. ISO 14044: 2006. Technical report, International Organization for Standardization, 2006.

[14] International Organization for Standardization. ISO standards - TC 207/SC 5 - life cycle assessment, 1 2013.

[15] E. Williams, C. Weber, and T. Hawkins. Hybrid framework for managing uncertainty in life cycle inventories. *Journal of Industrial Ecology*, 13(6):928–944, 2009.

[16] C. Weber, D. Matthews, A. Venkatesh, C. Costello, and H. Matthews. The 2002 US benchmark version of the economic input-output life cycle assessment (EIO-LCA) model. Technical report, The Green Design Institute of Carnegie Mellon University, 2010.

[17] K. Weitz, A. Sharma, B. Vigon, E. Price, G. Norris, P. Eagan, W. Owens, and A. Veroutis. Streamlined life-cycle assessment: A final report from the SETAC North America streamlined LCA workgroup. Society of Environmental Toxicology and Chemistry, SETAC Foundation for Environmental Education, July 1999.

[18] R. Hunt, T. Boguski, K. Weitz, and A. Sharma. Case studies examining LCA streamlining techniques. *The International Journal of Life Cycle Assessment*, 3(1):36–42, January 1998.

[19] T. Graedel. *Streamlined life-cycle assessment*. Prentice Hall, Upper Saddle River, NJ, 1998.

[20] T. Hur, J. Lee, J. Ryu, and E. Kwon. Simplified LCA and matrix methods in identifying the environmental aspects of a product system. *Journal of Environmental Management*, 75(3):229–237, 2005.

[21] S. Patanavanich, E. Olivetti, J. Gregory, and R. Kirchain. Exploring the viability of probabilistic underspecification to streamlined life cycle assessment, 2011.

[22] S. Patanavanich. Exploring the viability of probabilistic underspecification as a viable streamlining method for LCA. Master's thesis, Massachusetts Institute of Technology, September 2011.

[23] M. Zgola. A triage approach to streamline environmental footprinting : a case study for liquid crystal displays. Master's thesis, Massachusetts Institute of Technology, 2011.

[24] I. Sousa, J. Eisenhard, and D. Wallace. Approximate life-cycle assessment of product concepts using learning systems. *Journal of Industrial Ecology*, 4(4):61–81, 2001.

[25] I. Sousa. *Approximate Life-Cycle Assessment of Product Concepts Using Learning Systems*. PhD thesis, Massachusetts Institute of Technology, 2002.

[26] I. Sousa and D. Wallace. Product classification to support approximate life-cycle assessment of design concepts. *Technological Forecasting and Social Change*, 73(3):228–249, 2006.

[27] G. Sonnemann, M. Schuhmacher, and F. Castells. Uncertainty assessment by a Monte Carlo simulation in a life cycle inventory of electricity produced by a waste incinerator. *Journal of Cleaner Production*, 11(3):279–292, 2003.

[28] A. Boustani. Remanufacturing and energy savings. Master's thesis, Massachusetts Institute of Technology, 2010.

[29] Center for Transportation Analysis. Weight category definitions from 49cfr565.6. Technical report, Oak Ridge National Laboratory, 2000.

[30] Environmental Protection Agency. National clean diesel campaign (NCDC), 2013.

[31] Freedonia Group. World diesel engines. Technical report, The Freedonia Group, 767 Beta Drive, Cleveland, OH, April 2009.

[32] IBISWorld. Heavy-duty diesel engine manufacturing in the US: Market research report. Technical report, IBISWorld, 2012.

[33] Ward's Automotive Group. Motor vehicle facts and figures 2012. Technical report, Ward's Automotive Group, 2012.

[34] M. Auer. MonteCarlito - tool for Monte Carlo simulation, 2012.

[35] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. ISBN 3-900051-07-0.

[36] G. Pujol. *Sensitivity R Package: Sensitivity Analysis*, 2008. R package version 1.4-0.

[37] V. Smith and G. Keoleian. The value of remanufactured engines: Life-cycle environmental and economic perspectives. *Journal of Industrial Ecology*, 8(1-2):193–221, 2004.

[38] M. Huijbregts, L. Rombouts, S. Hellweg, R. Frischknecht, A. Hendriks, D. van de Meent, A. Ragas, L. Reijnders, and J. Struijs. Is cumulative fossil energy demand a useful indicator for the environmental performance of products? *Environmental Science & Technology*, 40(3):641–648, 2006.

[39] Ecoinvent. Ecoinvent database 2.2. Technical report, Swiss Centre for Life Cycle Inventories, 2010.

[40] R. Frischknecht, N. Jungbluth, H. J. Althaus, G. Doka, T. Heck, R. Hischier S. Hellweg, T. Nemecek, G. Rebitzer, M. Spielmann, and G. Wernet. Overview and methodology, ecoinvent report no. 1. Technical report, Swiss Centre for Life Cycle Inventories, 2007.

[41] G. Keoleian, G. Lewis, R. Coulon, V. Camobreco, and H. Teulon. LCI modeling challenges and solutions for a complex product system: A mid-sized automobile. *SAE transactions*, 107(6):1952–1963, 1998.

[42] The National Academies. Technologies and approaches to reducing the fuel consumption of medium- and heavy-duty vehicles. Technical report, The National Academies, 2010.

[43] G. Hammond and C. Jones. Embodied energy and carbon in construction materials. *Proceedings of the Institution of Civil Engineers - Energy*, 161(2):87–98, 2008.

[44] J. Li, P. Shrivastava, Z. Gao, and H. Zhang. Printed circuit board recycling: A state-of-the-art survey. *IEEE Transactions on Electronics Packaging Manufacturing*, 27(1):33–43, January 2004.

[45] R. Heijungs and R. Frischknecht. Representing statistical distributions for uncertain parameters in LCA. relationships between mathematical forms, their representation in ecospold, and their representation in CMLCA. *The International Journal of Life Cycle Assessment*, 10(4):248–254, 2005.

[46] M. Morgan and M. Henrion. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, June 1990.

[47] J. René van Dorp and S. Kotz. A novel extension of the triangular distribution and its parameter estimation. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 51(1):63–79, 2002.

[48] R. Heijungs and M. Huijbregts. A review of approaches to treat uncertainty in LCA. 2004.

[49] J. Cook. Numerical computation of stochastic inequality probabilities. Technical report, UT MD Anderson Cancer Center Department of Biostatistics Working Paper Series, 2008.

[50] J. Cook. Exact calculation of inequality probabilities. Technical report, UT MD Anderson Cancer Center Department of Biostatistics Working Paper Series, 2009.

[51] M. Wilk and R. Gnanadesikan. Probability plotting methods for the analysis of data. *Biometrika*, 55(1):1–17, 1968.

[52] B. Efron. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1):1–26, January 1979.

[53] M. Kendall. Partial rank correlation. *Biometrica*, 32(3):277–283, April 1942.

[54] R. Iman and J. Helton. An investigation of uncertainty and sensitivity analysis techniques for computer models. *Risk Analysis*, 8(1):71–90, 1988.

[55] J. Helton. Sampling-based methods for uncertainty and sensitivity analysis. In *Second Summer School on Sensitivity Analysis of Model Output*. University of Venice, June 2002.

[56] Oracle. Oracle® Crystal Ball: Sensitivity Chart Views, 2012.

[57] G. Geisler, S. Hellweg, and K. Hungerbuhler. Uncertainty analysis in life cycle assessment (LCA): Case study on plant-protection products and implications for decision making. *The International Journal of Life Cycle Assessment*, 10(3):192, 2005.

[58] Intergovernmental Panel on Climate Change. Glossary of the ipcc fourth assessment report. Technical report, Intergovernmental Panel on Climate Change, 2007.

[59] Green Building Challenge. Handbook of the green building challenge of the D-A-C-H clay brick industry. 2013.

[60] M. Huijbregts. Life-cycle impact assessment of acidifying and eutrophying air pollutants - claulation of equivalency factors with RAINS-LCA. Technical report, University of Amsterdam, 1999.

[61] Institute for Environment and Sustainability. Life cycle thinking. Technical report, European Commission - Join Research Centre, 2010.

[62] A. Azapagic, S. Perdan, and R. Clift. *Sustainable development in practice*. Wiley, 2004.

[63] K. Flury, N. Jungblutha, R. Frischknecht, and I. Muñoz. Recommendation for life cycle inventory analysis for water use and consumption. Working paper, ESU-services Ltd, 2012.

[64] International Energy Agency. Energy balances of OECD countries (2005-2006). Technical report, International Energy Agency, 2007.

[65] M. Branham. Semiconductors and sustainability: Energy and materials use in integrated circuit manufacturing. Master's thesis, Massachusetts Institute of Technology, 2008.

[66] T. Gutowski, C. Murphy, D. Allen, D. Bauer, B. Bras, T. Piwonka, P. Sheng, J. Sutherland, D. Thurston, and E. Wolff. Environmentally benign manufacturing: observations from Japan, Europe and the United States. *Journal of Cleaner Production*, 13(1):1–17, 2005.

[67] National Weather Service. NowData - NOAA weather data. Technical report, National Oceanic and Atmospheric Administration, 2013.

[68] Environmental Protection Agency. Introduction of cleaner-burning diesel fuel enables advanced pollution control for cars, trucks and buses. Technical report, Environmental Protection Agency, 2006.

[69] D. Espinoza, M. Goycoolea, E. Moreno, and A. Newman. Minelib: a library of open pit mining problems. *Annals of Operations Research*, pages 1–22, 2012.

[70] C. Gupta. *Extractive metallurgy of molybdenum.* CRC Press, 1992.

[71] J. Jorgenson. *Minerals Yearbook: Metals and Minerals 2006*, volume 1. US Geological Survey & Orienteering, 2009.

[72] Central Steel & Wire Company. 2006 Central Steel & Wire Company catalog. Chicago, IL, USA, 2006.

Appendix A

Analysis of Enterprise LCA Software

A.1 Introduction

As part of this research project, an enterprise-level LCA software was piloted for possible use by the engine manufacturer. The software chosen was Windchill LCA from PTC. This module integrates with PTC's product life cycle management (PLM) software, Windchill, to calculate the environmental impact of products. This software was installed on a remote server to simulate a production environment. The version tested was Windchill Product Analytics 10.1 M010 with the LCA Module.

As tested, the software included two databases for environmental impact data: Ecoinvent, version 2.2, released in 2010 by the Swiss Centre for Life Cycle Inventories [39] and the 2002 US Benchmark Version of the Economic Input-Output Life Cycle Assessment (EIO-LCA) Model with producer price, released in 2009 by the Green Design Institute at Carnegie Mellon University [16]. For this test only the Ecoinvent data was used. Component costs and the EIO-LCA model were not factored into the analysis. The categories in EIO-LCA are, in general, more broad and offer less resolution.

A.2 Applications

Windchill LCA was used for two distinct projects:

1. Analysis of a diesel engine's bill of materials (BOM), which includes the constituent parts, as well as the general processing of those parts.

2. Analysis of two different camshaft designs, including the detailed processing steps involved.

These two projects were chosen to test the usefulness of the tool in a broad range of applications from general (represented by the engine BOM analysis) to specific (represented by the camshaft analysis).

A.3 Material Processing

The primary focus of the Windchill LCA tool is material data. Therefore, special attention was given to testing its capabilities for analyzing processing. As tested, the module only has the ability to add materials to a part, not processes. Processes must instead be added as materials. Processes were divided into several general categories, which are described below.

The following process descriptions rely on scaling. Environmental impact as calculated in the Ecoinvent database scales linearly with process metrics, such as mass or surface area.

A.3.1 Processes that scale with mass

These processes scale with mass of the material they are applied to. Some act on an entire part, and others on a subset of the part.

A.3.2 Processes acting on the whole part

These processes act on the entire mass of the part. For this analysis, the mass of the part refers to the final mass of the part. One example of this type of process is heat treatment. To enter one of these processes into Windchill LCA, a separate "dummy" material must be created to represent each process. The mass of this material is then equal to that of the part.

A.3.3 Processing acting on a subset of the part

These processes also scale with mass, but act only on a subset of the part. All material removal processes, for example milling and drilling, are categorized as such. These processes can be added in a similar fashion as other mass-scaling processes. In some scenarios, listed at the end of this appendix, subset processes must be treated differently than processes acting on the whole part.

A.3.4 Processes that do not scale with mass

Many processes do not scale with mass, but instead with another metric. Most commonly this is surface area. Coating and plating are examples of such processes. These processes must be added with an artificial scaling factor, described in the Strategy section of this appendix.

A.3.5 Multi-part processes

Some processes are applied to multiple parts; the primary examples of these processes are painting and brazing. These processes may scale with mass, surface area, or another metric. These processes require a workaround, described in the Strategy section on the next page.

A.4 Assessment

A.4.1 Database

The included Ecoinvent database is extremely extensive and covers many materials and processes. In our testing, this information was very helpful and expedited our work. Conversely, the database does have a limited amount of data on certain materials. For example, titanium and tool steel are not included. The processes are also relatively general, and there is little ability to differentiate between processes. As an example, the database only includes one general heat treatment process for steel, as opposed to specific processes such as carburizing or ammonia gas nitriding.

A.4.2 Indicators

The Ecoinvent tool provides many different ecological indicators. The Windchill LCA tool provides access to five of the most common: greenhouse gas emissions, energy use, water use, acidification potential, and eutrophication potential. These five indicators should suffice for most private and public uses.

A.4.3 Features and Stability

The LCA module was tested after its first release. A variety of bugs were encountered during the testing phase. The feature set was also limited in some areas. Various workarounds are essential in the module's current state. For example, the module cannot utilize processing data without implementing a workaround. Compromises must therefore be made between conflicting priorities. The following section, Strategies, serves as a brief guide to the major decisions.

A.5 Strategies

To overcome the aforementioned shortcomings of the tool, a variety of workarounds may be employed. Listed below are several strategies, or general approaches to the workarounds and compromises. The strategies are listed roughly in order of increasing effort. Each strategy has both benefits and disadvantages. There is currently not an ideal strategy, regardless of effort expended.

The advantages and disadvantages associated with all strategies discussed below are summarized in Table A.1.

A.5.1 Default Strategy

In the default strategy, the software is used as provided. Material information is included for every part; however, no process data is included. Calculated part mass is the sum of the material masses; as such, the transportation calculations, which scale with part mass, are valid.

A.5.2 Strategy A: Basic Processes

This strategy modifies the default strategy by including mass-based processes as "dummy" materials. Each such new material must be created, and therefore this strategy requires moderately more effort. The primary disadvantage of this strategy is that the transportation calculations are no longer valid, as the calculated mass of the part is all of the materials plus each process. For example, if a part has 10% of its mass removed by machining and then undergoes heat treatment, the calculated mass is the sum of materials, machining, and heat treatment: 210% of the actual part mass.

A.5.3 Strategy B: Customized Processes

Strategy B builds on A by adding the remaining processes that do not scale with mass. All processes are now included, each as a separate dummy material. This process requires a moderate effort.

The primary disadvantage of this strategy is that the units of the new processes are not accurate. As the tool only accepts materials in terms of mass, the new dummy materials must have an artificial conversion factor. For example, if a part is painted, this new process will scale with area, measured in m2. The units of the dummy material will be in mass, kg; there must therefore be an arbitrary and artificial conversion from m2 to kg. For these tests, the conversion was a 1-to-1 conversion from m2 to kg.

Strategy B suffers from the same invalid transportation calculations as A. The transportation impact calculations are therefore more skewed due to the additional processes and artificial units.

A.5.4 Strategy C: Lumped processes and materials

This strategy overcomes the invalid transportation calculations previously mentioned. Many additional dummy materials are created which include both material and processing information. The calculated mass of the part is therefore accurate, and the transportation impact calculations are valid.

This strategy requires much more effort than A or B, as the number of custom materials will be larger than the total number of materials. For example, aluminum with and without heat treatment must be created as two separate materials. If a second process option, such as coating, is added, the number of custom materials is now four. Any additional process options increase the number of custom materials exponentially.

Strategy C has significant drawbacks. Because processes and materials are combined, any process that does not scale with the final mass of the part cannot be included. This then excludes processes such as machining and coating.

A.5.5 Strategy D: Customized processes with unit scaling

Strategy D is a simple but consequential variation of B. The advantage of this approach is the validity of the transportation calculations. The units of all processes, without exception, are scaled significantly. The objective is to reduce the added mass from a process to such a degree that it becomes insignificant in part mass calculations.

For example, the impact of all mass-based processes may increase by 1,000. Therefore, removing 2 kg of material by machining would be entered into the software as 0.002 kg of machining. The conversion to the appropriate 2 kg of machining must be included in the custom material for machining. The primary and significant disadvantage of this strategy is the misleading artificial units of all the processing steps.

A.5.6 Strategy E: Custom material for each part

This strategy is the most time-intensive and offers no scalability. Strategy E is identical to C, except that each part has a unique custom material. This enables valid transportation calculations and any processing, but no reusable materials or processes.

A.5.7 Multi-part processes

Multi-part processes may be added through the addition of dummy parts (as opposed to dummy materials). These parts could then have the appropriate processes attached to them. The primary disadvantage of this approach is that the BOM must be modified, which may not be appropriate for a production environment.

Attributes	Strategies						
	∅	*A*	*B*	*C*	*D*	*F*	*E*
Valid transportation results	●			●	●	●	●
Materials	●	●	●	●	●	●	●
Mass-based processes, whole part		●	●	●	○	●	●
Other mass-based processes		●	●		○	○	●
Other processes			○		○	○	●
Multi-part processes						○	○
BOM used as-is	●	●	●	●	●		
Relative Complexity	0	1	2	3	3	3	4

Table A.1 – Possible strategies and corresponding attributes when implementing Windchill LCA. The default implementation is represented as $\emptyset$. Strategies A through E are described in the text above. The ● symbol signifies that an attribute is available or true. The ○ symbol applies only to process data, and signifies that the processes are available, but the associated units are incorrect. These processes are then usable and the calculated results correct; however, interpretation and scalability suffers. The Relative Complexity attribute is a qualitative measure of the effort and complexity of the strategy; the lower the complexity, the easier and more straightforward a strategy.

A.6 Primary Shortcomings

The following were presented to the software maker PTC as the primary shortcomings of the tool.

1. No material processing support. There was no native functionality to add process information without artificially adding to the mass of the part.

2. Non-mass-based processes. There was no native support for any processing steps that did not scale with mass.

Appendix B

Additional Data Tables, Figures, and Equations

Material	Lower ($^{MJ}/_{kg}$)	Upper ($^{MJ}/_{kg}$)
Aluminum Alloy	44	59
Copper	25	40
Copper-Tin Alloy	60	76
Gray Iron	18	29
Low Alloy Steel	20	34
Nickel	158	240
Nickel Alloy	132	205
Nylon 6-6	86	128
PCB, General	9,963	14,239
PCB, Logic Type	12,579	20,209
PCB, Memory Type	8,240	11,708
PET Plastic	67	102
Stainless Steel	60	89
Tin	233	384
Titanium	605	833
Turbocharger Alloy	36	62

Table B.1 – Material energy intensity ranges. These ranges form the initial rough estimate for material energy intensity of the engine componentes. Through resolution these values can change independently for each part. This data aggregated from Ecoinvent [39], Ashby [9], and Hammond and Jones [43].

Mode	Lower $^e/_d$, $^{MJ}/_{t\cdot km}$	Upper $^e/_d$, $^{MJ}/_{t\cdot km}$
Truck	1.7 *transport, lorry >32t, EURO5*	7.3 *transport, lorry 3.5-7.5t, EURO3*
Rail	0.49 *transport, freight, rail, Austria*	0.70 *transport, freight, rail, diesel, US*
Ship	0.15 *transport, transoceanic freight ship*	0.61 *transport, barge*

Table B.2 – Range of transportation mode energy intensities and their associated source in the Ecoinvent database. Note that these values are given per metric ton rather than kg.

$\lvert PRCC \rvert$ Rank	Category	Description	$\lvert PRCC \rvert$	Resolvable
1	Material Composition	Remaining aluminum fraction	0.738	Yes
2	Process Mass	Machining of remaining steel	0.717	-
3	Material Intensity	Remaining low alloy steel	0.649	-
4	Process Intensity	Machining of cylinder block	0.630	Yes
5	Process Intensity	Forging of remaining steel	0.533	-
6	Transit Distance	Remaining mass	0.500	-
7	Material Composition	Remaining gray iron fraction	0.498	Yes
8	Material Intensity	Cylinder block, gray iron	0.497	Yes
9	Process Intensity	Machining of remaining aluminum	0.492	-
10	Process Intensity	Forging of engine crankshaft	0.482	Yes
11	Material Composition	Remaining titanium fraction	0.473	Yes
12	Process Intensity	Machining of cylinder head	0.468	Yes
13	Transit Mode	Remaining transportation, truck	0.425	-
14	Part Mass	Electronic control module	0.415	Yes
15	Process Intensity	Machining of crankshaft	0.391	Yes
16	Material Intensity	Crankshaft, low alloy steel	0.373	Yes
17	Material Intensity	Cylinder head, gray iron	0.355	Yes
18	Process Intensity	Machining of remaining steel	0.328	-
19	Transit Mode	Remaining transportation, ship	0.249	-
20	Process Intensity	Machining of cylinder block	0.215	Yes
21	Transit Mode	Remaining transportation, rail	0.208	-
22	Material Composition	Remaining tin fraction	0.198	-
23	Transit Intensity	Remaining transportation, truck	0.196	-
24	Material Intensity	Electronic control module, PCB	0.177	Yes
25	Part Mass	Total engine mass	0.159	Yes

Table B.3 – Parameters with the 25 largest $\lvert PRCC \rvert$ values after 1^{st} round of simulations.

Table B.4 – False Signal Rate details after all simulation rounds.

Parameters Resolved	Parameter Category	Parameter Description	FSR
0	-	Initial Round	21.5%
1	Part Mass	Total engine mass	11.9%
2	Material Composition	Remaining aluminum fraction	9.7%
3	Material Composition	Remaining gray iron fraction	9.6%
4	Process Mass	Machining of cylinder block	9.4%
5	Material Intensity	Cylinder block, gray iron	8.4%
6	Material Composition	Remaining titanium fraction	8.4%
7	Process Intensity	Forging of engine crankshaft	7.7%
8	Process Mass	Machining of cylinder head	7.6%
9	Part Mass	Electronic control module	7.3%
10	Process Mass	Machining of crankshaft	6.7%
11	Material Intensity	Crankshaft, low alloy steel	6.4%
12	Material Intensity	Cylinder head, gray iron	6.1%
13	Process Intensity	Machining of cylinder block	6.1%
14	Material Composition	Remaining tin fraction	6.1%
15	Material Intensity	Electronic control module, PCB	6.1%
16	Process Intensity	Machining of cylinder head	6.0%
17	Process Mass	Machining of turbocharger	6.0%
18	Process Intensity	Heat treatment of cylinder block	6.3%
19	Process Intensity	Machining of crankshaft	6.0%

continued...

Table B.4: (continued)

Parameters Resolved	Parameter Category	Parameter Description	FSR
20	Material Intensity	Turbocharger, Al and Ni	6.1%
21	Process Intensity	Machining of oil filter head	5.8%
22	Process Intensity	Coating of air intake manifold	5.8%
23	Material Composition	Remaining stainless steel fraction	6.2%
24	Process Intensity	Heat treatment of cylinder head	6.4%
25	Part Mass	Oil filter head, aluminum	6.4%
26	Process Mass	Machining of air intake manifold	6.3%
27	Part Mass	Air intake manifold	6.2%
28	Process Intensity	Heat treatment of crankshaft	6.0%
29	Process Mass	Machining of flywheel	6.4%
34	Various	Group A[1]	6.4%
39	Various	Group B[1]	5.6%
44	Various	Group C[1]	5.5%
49	Various	Group D[1]	5.5%
76	Part & Process Mass	Remaining 27 parameters	5.4%
101	Material Intensity	Remaining 25 parameters	5.6%
130	Process Intensity	Remaining 29 parameters	5.6%
133	Transit Distance	Remaining 3 parameters	5.5%
142	Transit Mode	Remaining 9 parameters	5.5%
151	Transit Intensity	Remaining 9 parameters	5.8%

These data points were taken after restarting the simulations. Note that the FSR increases after some rounds. This is explored more in Table B.7 on page 121.

[1]See Table B.6 on page 119, which describes the parameters placed in these groups.

Table B.5 – Leverage† of each parameter for uncertainty reduction.

Parameters Resolved	Parameter Category	Parameter Description	Leverage†
1	Part Mass	Total engine mass	●●
2	Material Composition	Remaining aluminum fraction	●●
3	Material Composition	Remaining gray iron fraction	○
4	Process Mass	Machining of cylinder block	●
5	Material Intensity	Cylinder block, gray iron	●●
6	Material Composition	Remaining titanium fraction	
7	Process Intensity	Forging of engine crankshaft	●●
8	Process Mass	Machining of cylinder head	●
9	Part Mass	Electronic control module	●
10	Process Mass	Machining of crankshaft	●●
11	Material Intensity	Crankshaft, low alloy steel	●
12	Material Intensity	Cylinder head, gray iron	●
13	Process Intensity	Machining of cylinder block	○
14	Material Composition	Remaining tin fraction	
15	Material Intensity	Electronic control module, PCB	
16	Process Intensity	Machining of cylinder head	●
17	Process Mass	Machining of turbocharger	
18	Process Intensity	Heat treatment of cylinder block	
19	Process Intensity	Machining of crankshaft	●

continued...

Table B.5: (continued)

Parameters Resolved	Parameter Category	Parameter Description	Leverage†
20	Material Intensity	Turbocharger, Al and Ni	
21	Process Intensity	Machining of oil filter head	●
22	Process Intensity	Coating of air intake manifold	
23	Material Composition	Remaining stainless steel fraction	
24	Process Intensity	Heat treatment of cylinder head	
25	Part Mass	Oil filter head, aluminum	
26	Process Mass	Machining of air intake manifold	○
27	Part Mass	Air intake manifold	●
28	Process Intensity	Heat treatment of crankshaft	●
29	Process Mass	Machining of flywheel	
34	Various	Group A[2]	○
39	Various	Group B[2]	●●
44	Various	Group C[2]	●
49	Various	Group D[2]	
76	Part & Process Mass	Remaining 27 parameters	○
101	Material Intensity	Remaining 25 parameters	
130	Process Intensity	Remaining 29 parameters	
133	Transit Distance	Remaining 3 parameters	●
142	Transit Mode	Remaining 9 parameters	○
151	Transit Intensity	Remaining 9 parameters	

†Parameter leverage over model uncertainty was judged by the median change in FSR from the previous round, as measured in the bootstrap resampling.

●● - Significant leverage. $|\Delta\text{FSR}| > 0.50\%$

● - Moderately leverage. $|\Delta\text{FSR}| > 0.10\%$

○ - Trivial leverage. $|\Delta\text{FSR}| > 0.03\%$

blank - Insignificant parameter. $|\Delta\text{FSR}| < 0.03\%$

[2]See Table B.6 on the facing page, which describes the parameters placed in these groups.

Group	Category	Description
Group A	Part Mass	Fuel Filter Head
	Process Intensity	Machining of Fuel Filter Head
	Part Mass	Oil Pan
	Material Intensity	Oil Pan, Low Alloy Steel
	Process Intensity	Machining of Gear Housing
Group B	Process Intensity	Machining of Fuel Pump Head
	Material Intensity	Flywheel, Gray Iron
	Part Mass	Cylinder Head, Gray Iron
	Material Intensity	Oil Filter Head
	Process Intensity	Forging of Oil Filter Head
Group C	Process Intensity	Machining of Oil Filter Head
	Part Mass	Cylinder Block
	Part Mass	Turbocharger
	Material Intensity	Valve Cover PET
	Process Intensity	Forging of Fuel Pump Head
Group D	Process Intensity	Machining of Turbocharger
	Part Mass	Engine Crankshaft
	Part Mass	Gear Housing Cast Aluminum Alloy
	Process Intensity	Machining of Camshaft
	Process Intensity	Forging of Connecting Rod

Table B.6 – Grouped parameters. After 30 rounds of simulations, the next 20 parameters to be resolved were placed into groups of 5 each and resolved together.

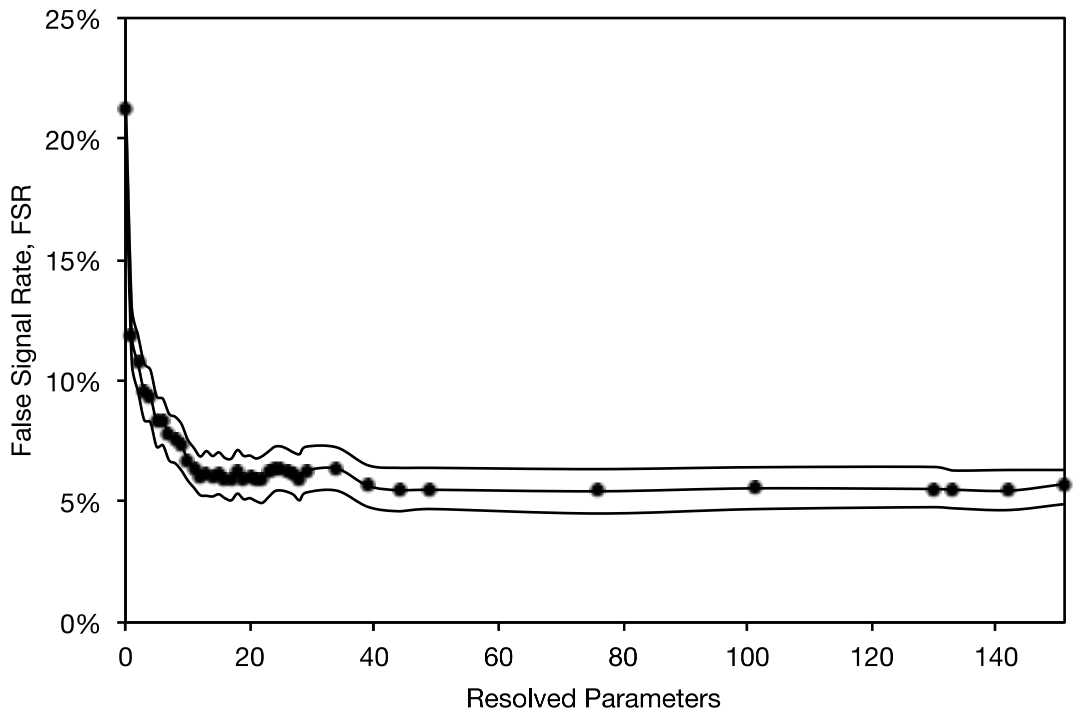


Figure B.1 – False Signal Rate for final rounds. Bootstrap resampling was used to produce the 5% and 95% quantiles shown on the figure. False Signal Rate for final rounds.

Resolved Parameters	FSR Bootstrap Percentile			
	5^{th}	50^{th}	60^{th}	95^{th}
0	20.08%	21.23%	21.58%	22.27%
1	10.76%	11.88%	12.24%	13.07%
2	9.57%	10.72%	11.10%	11.89%
3	8.37%	9.56%	9.96%	10.72%
4	8.28%	9.33%	9.71%	10.44%
5	7.27%	8.32%	8.72%	9.35%
6	7.33%	8.36%	8.65%	9.26%
7	6.71%	7.74%	8.02%	8.63%
8	6.58%	7.61%	7.90%	8.50%
9	6.28%	7.31%	7.60%	8.20%
10	5.87%	6.73%	7.01%	7.56%
11	5.57%	6.40%	6.66%	7.21%
12	5.27%	6.06%	6.32%	6.86%
13	5.24%	6.13%	6.38%	7.09%
14	5.21%	6.08%	6.38%	6.88%
15	5.30%	6.11%	6.40%	7.04%
16	5.14%	5.94%	6.23%	6.80%
17	5.05%	5.95%	6.23%	6.77%
18	5.33%	6.23%	6.52%	7.13%
19	5.10%	5.95%	6.26%	6.89%
20	5.14%	6.05%	6.30%	6.90%
21	5.02%	5.88%	6.18%	6.78%
22	4.95%	5.87%	6.21%	6.88%
23	5.19%	**6.25%†**	6.42%	7.07%
24	5.43%	6.31%	6.63%	7.27%
25	5.46%	6.32%	6.61%	7.28%
26	5.40%	6.24%	6.52%	7.18%
27	5.29%	6.18%	6.43%	7.03%
28	5.05%	5.95%	6.25%	6.96%
29	5.37%	**6.28%†**	6.63%	7.27%
34	5.45%	6.36%	6.61%	7.24%
39	4.78%	5.65%	5.90%	6.50%
44	4.60%	5.47%	5.76%	6.39%
49	4.69%	5.49%	5.76%	6.40%
76	4.51%	5.42%	5.67%	6.34%
101	4.69%	5.55%	5.81%	6.42%
130	4.77%	5.52%	5.83%	6.43%
133	4.72%	5.49%	5.76%	6.29%
142	4.65%	5.45%	5.71%	6.31%
151	4.90%	**5.73%†**	5.99%	6.61%

Table B.7 – Selected percentiles of the False Signal Rate of resampled distributions. Only 3 rounds have an FSR greater than the 60^{th} percentile of the previous round's FSR; these are set in bold and marked with a cross (†): 23, 29, and 151.

Part	Embodied Energy (GJ)	Greenhouse Gas Emissions (ton CO_2-eq)	Acidification Potential (kg SO_2-eq)	Eutrophication Potential (kg PO_4-eq)	Embodied Water Use (kl)
Cylinder Block	13.5	0.82	4.4	1.4	4.6
Engine Crankshaft	10.9	0.63	2.5	1.5	8.4
Cylinder Head	8.3	0.51	2.0	0.88	3.4
Electronic Control Module	2.2	0.13	0.65	1.4	1.4
Turbocharger	1.6	0.10	3.4	0.62	1.2
Connecting Rod	1.4	0.08	0.33	0.22	1.2
Piston	1.2	0.07	0.35	0.13	0.54
Oil Filter Head	0.65	0.04	0.17	0.09	0.36
Air Intake Manifold	0.61	0.04	0.21	0.07	0.33
Fuel Pump Head	0.51	0.03	0.11	0.07	0.43
Oil Pan	0.43	0.03	0.10	0.06	0.44
Flywheel	0.40	0.03	0.10	0.05	0.24
Gear Housing	0.37	0.02	0.10	0.05	0.21
Exhaust Transfer Tube	0.36	0.02	0.12	0.04	0.16
Fuel Filter Head	0.32	0.02	0.08	0.04	0.18
Camshaft	0.32	0.02	0.07	0.04	0.32
Valve Cover	0.25	0.01	0.03	0.01	0.05
Exhaust Tube	0.23	0.01	0.07	0.03	0.10
Gear Cover	0.22	0.01	0.07	0.04	0.17
Flywheel Housing	0.20	0.01	0.05	0.03	0.16
Exhaust Manifold	0.19	0.01	0.04	0.02	0.10
Fuel Pump Housing	0.18	0.01	0.05	0.02	0.06
Exhaust Manifold	0.17	0.01	0.04	0.02	0.10
Accumulator	0.15	0.01	0.04	0.02	0.12
Nitrogen Sensor	0.11	0.01	0.19	0.12	0.34
All Other Parts	26	1.6	17.0	5.6	18.3

Table B.8 – Results from streamlined LCA, by part. The most energy-intensive 25 parts are included. The same results are presented as fractional values in Table B.9 on the facing page.

Part	Embodied Energy	Greenhouse Gas Emissions	Acidification Potential	Eutrophication Potential	Embodied Water Use
Cylinder Block	19.0%	19.0%	13.5%	11.4%	10.7%
Engine Crankshaft	15.3%	14.6%	7.8%	12.2%	19.6%
Cylinder Head	11.7%	11.7%	6.3%	7.0%	8.0%
Electronic Control Module	3.1%	3.0%	2.0%	11.0%	3.3%
Turbocharger	2.3%	2.2%	10.6%	4.9%	2.8%
Connecting Rod	2.0%	2.0%	1.0%	1.7%	2.8%
Piston	1.6%	1.6%	1.1%	1.0%	1.3%
Oil Filter Head	0.9%	0.9%	0.5%	0.7%	0.8%
Air Intake Manifold	0.9%	1.0%	0.7%	0.5%	0.8%
Fuel Pump Head	0.7%	0.7%	0.4%	0.6%	1.0%
Oil Pan	0.6%	0.7%	0.3%	0.5%	1.0%
Flywheel	0.6%	0.6%	0.3%	0.4%	0.5%
Gear Housing	0.5%	0.5%	0.3%	0.4%	0.5%
Exhaust Transfer Tube	0.5%	0.5%	0.4%	0.3%	0.4%
Fuel Filter Head	0.5%	0.4%	0.3%	0.3%	0.4%
Camshaft	0.4%	0.5%	0.2%	0.3%	0.7%
Valve Cover	0.4%	0.2%	0.1%	0.1%	0.1%
Exhaust Tube	0.3%	0.3%	0.2%	0.2%	0.2%
Gear Cover	0.3%	0.3%	0.2%	0.3%	0.4%
Flywheel Housing	0.3%	0.3%	0.1%	0.2%	0.4%
Exhaust Manifold	0.3%	0.3%	0.1%	0.2%	0.2%
Fuel Pump Housing	0.3%	0.3%	0.2%	0.2%	0.1%
Exhaust Manifold	0.2%	0.3%	0.1%	0.2%	0.2%
Accumulator	0.2%	0.2%	0.1%	0.2%	0.3%
Nitrogen Sensor	0.1%	0.1%	0.6%	0.9%	0.8%
All Other Parts	37.0%	37.8%	52.6%	44.4%	42.6%

Table B.9 – Results from streamlined LCA, by part. All part impacts are given as a fraction of the total engine impact for that metric. The most energy-intensive 25 parts are included.

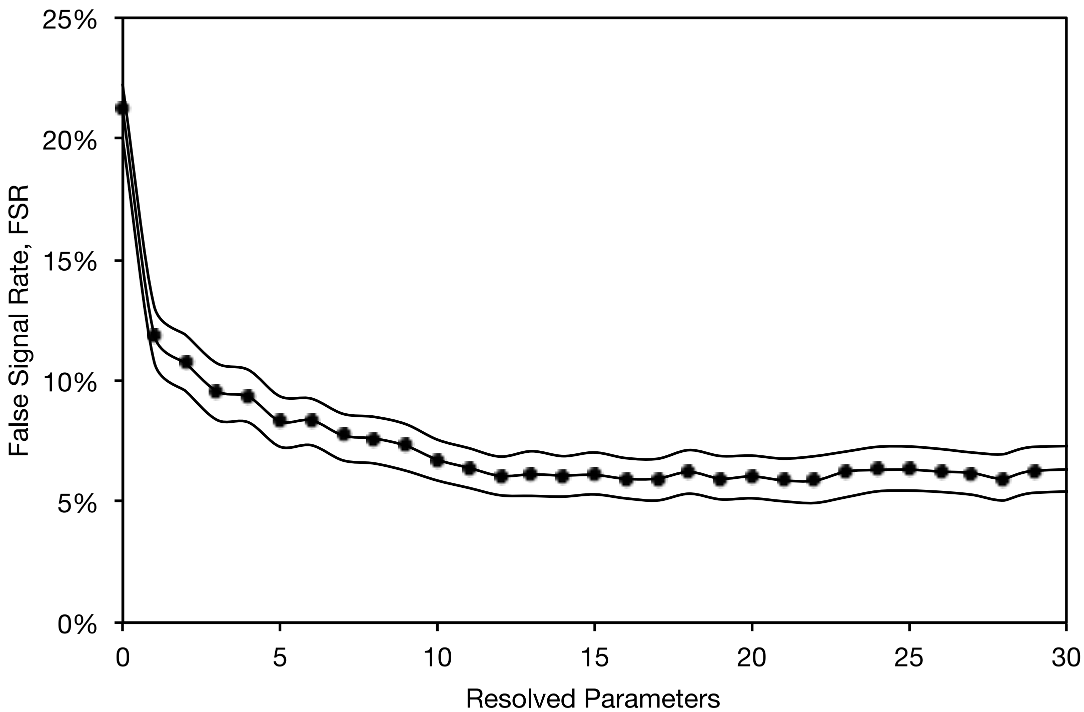


Figure B.2 – False Signal Rate for final rounds (first 30). Bootstrap resampling was used to produce the 5% and 95% quantiles shown on the figure.

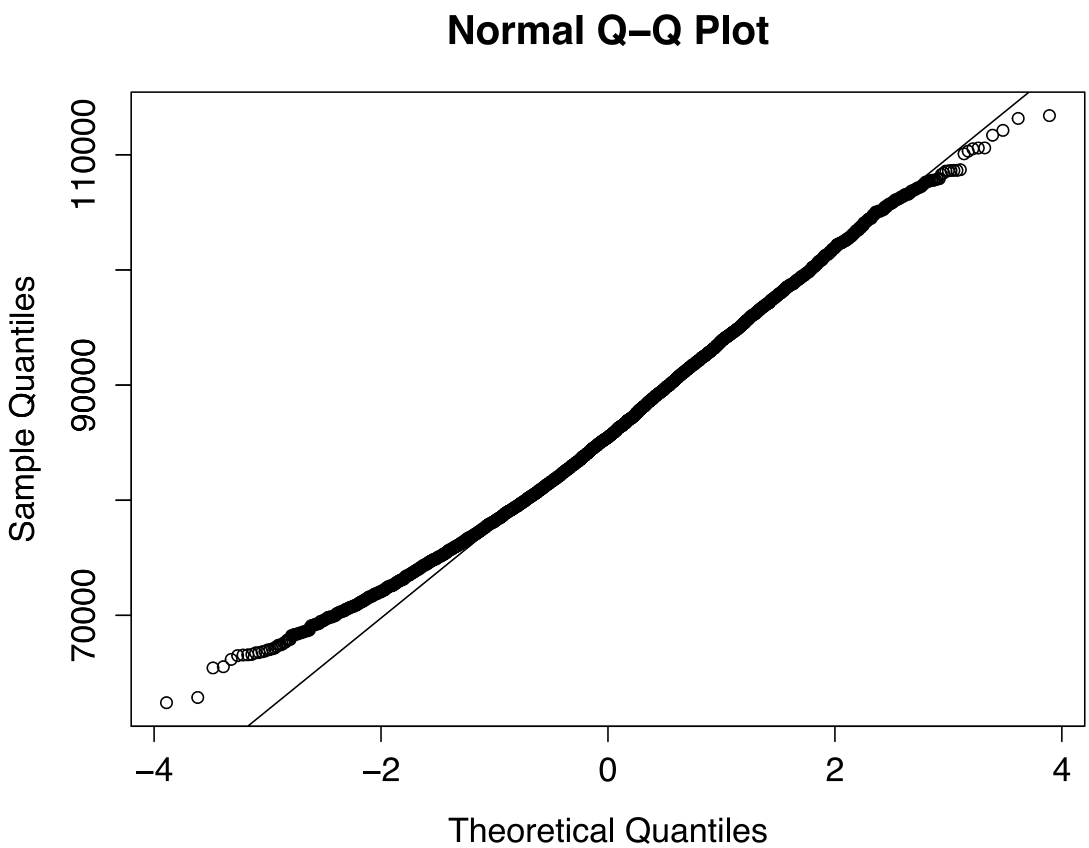


Figure B.3 – Normal Q-Q plot (quantile-quantile plot) for the distribution of estimated engine impact. Data from the model with no parameters resolved. The normal Q-Q plot is useful when arguing qualitatively for the normality of a specific distribution. From this plot it can be inferred that the impact distribution is indeed approximately Gaussian. See Wilk [51] for an introduction to Q-Q plots.

Table B.10 – Embodied energy for selected materials from the diesel engine case study. Singular representative values are shown in this table for simplicity; however, it must be noted that each material has a broad range of embodied energy values.

Material	Embodied Energy ($^{MJ}/_{kg}$)	Data Source
Aluminum Alloy	72	Ecoinvent, index #1045, *"aluminium alloy, AlMg3, at plant"*
Aluminum	51	Ecoinvent, index #1057, *"aluminium, production mix, cast alloy, at plant"*
Copper	34	Ecoinvent, index #1074, *"copper, at regional storage"*
Steel, Cold Rolled Coil	28	Ecoinvent, index #1154, *"steel, low-alloyed, at plant"*
Ferromanganese	23	Ecoinvent, index #1097, *"ferromanganese, high-coal, 74.5% Mn, at regional storage"*
Ferromolybdenum	43	Composite (See Gutpa [70]). 65% molybdenum: Ecoinvent, index #1118, *"Molybdenum concentrate, main product."* 35% pig iron: Ecoinvent, index #1132, *"pig iron, at plant."*
Ferrosilicon (Fe-Si)	16	Composite (See [71]). 45% low-grade silicon: Ecoinvent, index #321, *"silicon carbide, at plant."* 55% pig iron: Ecoinvent, index #1132, *"pig iron, at plant."*
Steel, Hot Rolled Coil	28	Ecoinvent, index #1154, *"steel, low-alloyed, at plant"*

continued...

Table B.10 on the preceding page: (continued)

Material	Embodied Energy (MJ/kg)	Data Source
Steel, 4140	33	Composite (See Central Steel & Wire Company [72]). 99% low-alloy steel: Ecoinvent, index #1154, *"steel, low-alloyed, at plant."* 1% chromium: Ecoinvent, index #1073, *"chromium, at regional storage."*
Molybdenum	151	Ecoinvent, index #1116, *"molybdenum, at regional storage"*
Nickel	142	Recycled production mix (See Ashby [9]). 74% primary nickel: Ecoinvent, index #1121 nickel, *"99.5%, at plant."* 26% Ecoinvent, index #8149, *"nickel, secondary, from electronic and electric scrap recycling, at refinery."*
Pig Iron	23	Ecoinvent, index #1132, *"pig iron, at plant"*
Iron Scrap	0.73	Ecoinvent, index #1101, *"iron scrap, at plant"*
Lead	16	Ecoinvent, index #1103, *"lead, at regional storage"*
Tin	321	Ecoinvent, index #1155, *"tin, at regional storage"*
Zinc	52	Ecoinvent, index #1156, *"zinc, primary, at regional storage"*
Titanium	670	Ashby [9], mean value from range of 600 - 740 MJ/kg.
Stainless Steel	68	Ecoinvent, index #1152, *"steel, electric, chromium steel 18/8, at plant"*

Impact variance demonstration equation, where i is impact, e is energy intensity, m is mass, and a and b are parts:

$$i = e_a m_a + e_b m_b \tag{B.1}$$

Definitions and givens for Equations B.9 through B.12:

$$
\begin{align}
I &= E_a M_a + E_b M_b \tag{B.2} \\
m_a + m_b &= c \tag{B.3} \\
M_a &= -M_b + c \tag{B.4} \\
\mathrm{Var}(M_a) &= \mathrm{Var}(-M_b) \tag{B.5} \\
&= (-1)^2 \mathrm{Var}(M_b) \nonumber \\
&= \mathrm{Var}(M_b) \nonumber \\
\mathrm{Cov}(E_a, E_b) &= 0 \tag{B.6} \\
\mathrm{Cov}(E_a, M_a) &= 0 \tag{B.7} \\
\mathrm{Cov}(E_b, M_b) &= 0 \tag{B.8}
\end{align}
$$

Variance of product of energy intensity and mass:

$$\mathrm{Var}(E_a M_a) = \mathrm{E}(\mathrm{E_a})^2 \cdot \mathrm{Var}(E_a) + \mathrm{E}(M_a)^2 \cdot \mathrm{Var}(M_a) + \mathrm{Var}(E_a) \cdot \mathrm{Var}(M_a) \tag{B.9}$$

Covariance of masses:

$$\begin{aligned}
\mathrm{Cov}(M_a, M_b) &= \mathrm{E}(M_a M_b) - \mathrm{E}(M_a)\mathrm{E}(M_b) \\
&= \mathrm{E}((-M_b + c) \cdot M_b) - \mathrm{E}(M_a)\mathrm{E}(M_b) \\
&= \mathrm{E}(-M_b^2 + cM_b) - \mathrm{E}(M_a)\mathrm{E}(M_b) \\
&= -\mathrm{E}(M_b^2) + c\mathrm{E}(M_b) - \mathrm{E}(M_a)\mathrm{E}(M_b) \\
&= -\mathrm{E}(M_b^2) + \mathrm{E}(M_b)(c - \mathrm{E}(M_a)) \\
&= -\mathrm{E}(M_b^2) + \mathrm{E}(M_b)(\mathrm{E}(c - M_a)) \\
&= -(\mathrm{E}(M_b^2) - \mathrm{E}(M_b)^2) \\
&= -\mathrm{Var}(M_b)
\end{aligned} \tag{B.10}$$

Covariance of $E_a M_a$ and $E_b M_b$:

$$\begin{aligned}
\mathrm{Cov}(E_a M_a, E_b M_b) &= \mathrm{E}(E_a M_a \cdot E_b M_b) - \mathrm{E}(E_a M_a) \cdot \mathrm{E}(E_b M_b) \\
&= \mathrm{E}(E_a M_a \cdot E_b M_b) - \mathrm{E}(E_a M_a) \cdot \mathrm{E}(E_b M_b) \\
&= \mathrm{E}(E_a) \cdot \mathrm{E}(E_b) \cdot \mathrm{E}(M_a M_b) - \mathrm{E}(E_a) \cdot \mathrm{E}(M_a) \cdot \mathrm{E}(E_b) \cdot \mathrm{E}(M_b) \\
&= \mathrm{E}(E_a) \cdot \mathrm{E}(E_b) \left(\mathrm{E}(M_a M_b) - \mathrm{E}(M_a)\mathrm{E}(M_b)\right) \\
&= \mathrm{E}(E_a) \cdot \mathrm{E}(E_b)\mathrm{Cov}(M_a, M_b)
\end{aligned} \tag{B.11}$$

Variance of impact:

$$\begin{aligned}
\mathrm{Var}(I) =& \mathrm{Var}(E_a M_a + E_b M_b) \\
=& \mathrm{Var}(E_a M_a) + \mathrm{Var}(E_b M_b) + 2\mathrm{Cov}(E_a M_a, E_b M_b) \\
=& (\mathrm{E}(\mathrm{E_a})^2 \cdot \mathrm{Var}(E_a) + \mathrm{E}(M_a)^2 \cdot \mathrm{Var}(M_a) + \mathrm{Var}(E_a) \cdot \mathrm{Var}(M_a)) \\
&+ (\mathrm{E}(\mathrm{E_b})^2 \cdot \mathrm{Var}(E_b) + \mathrm{E}(M_b)^2 \cdot \mathrm{Var}(M_b) + \mathrm{Var}(E_b) \cdot \mathrm{Var}(M_b)) \\
&- 2(\mathrm{E}(E_a) \cdot \mathrm{E}(E_b) \cdot \mathrm{Var}(M_b))
\end{aligned} \tag{B.12}$$